lake.) This illogical conclusion is the result of the bid design, not the true preferences of Oklahoma residents. Of course, these interpretations require one to ignore the confounding effects we discussed in the previous section.[51]

4.6 The Stratus survey contains nonresponse bias.

The response rate reflects the portion of the intended sample that actually participates in the study. The consequence of a low response rate is nonresponse bias in the data. This bias occurs when respondents to the survey are systematically different from those who do not respond. The most serious concern about nonresponse bias is that there are likely to be unknown and in fact unknowable differences between nonrespondents and the people who completed the survey.[52] As the empirical analysis described in Section 5 demonstrates, the WTP models vary substantially with differences in individual characteristics, attitudes, and experiences. To merely make adjustments for differences in a few demographic characteristics as the Stratus report does, is not to account for the most serious consequences of nonresponse bias. As a result, the data collected do not accurately reflect the responses of the target population. CV surveys with a high nonresponse rate (or a low response rate) are considered "unreliable" by the NOAA Panel (Arrow, et al. 1993).

The response rate is a critical indicator of the quality of the data from the study. For the Stratus study, the response rate ranged between 52 to 57 percent depending on various assumptions of eligibility and other survey features. According to Smith (2007), the NOAA Panel defined 70 percent as a high response rate. The response rate of a similar CV study conducted for NRD purposes was 72.6 percent (Carson, et al. 1994). As the Stratus report indicates, the Office of Management and Budget (OMB) guidelines for conducting surveys (2006) establish a threshold response rate of

[51] Additional evidence supporting the perspective presented in this section include Carson, et al. 1996 and Dunford, et al. 1996. For example, in Carson, et al. (1996) dropping the highest bid offered ($220) lowers the mean from $85 to $56, or a decrease of 34 percent. Dunford, et al. (1996) re-estimate the COS mean using alternative bid structures. They find that adding a higher bid of $400 increases the mean to $124, an increase of 48 percent. This degree of sensitivity to the bid structure indicates an overall lack of reliability of the CV method, especially because the selection of bids is essentially arbitrary, resting solely under the control of the survey designers.

[52] Smith (2007) discusses the problem as one of unobserved heterogeneity. That is, respondents and nonrespondents differ in ways that cannot be measured leading to un-interpretable results.

80 percent. The response rate in the Stratus CV study is substantially below all of these benchmarks.

In an attempt to address nonresponse bias in its CV data, Stratus uses available demographic information to re-weight the data. In addition, Stratus implements only two of the several analytical assessments recommended by OMB for surveys with response rates below 80 percent. First, Stratus compares the age, gender, race/ethnicity, and educational distributions of the survey respondents to Census data for Oklahoma. From this analysis, Stratus concludes that the survey respondents mirror the State residents with respect to these four demographic characteristics. Second, Stratus compares the percentage of "for" votes for the early and late respondents and the percentage of "for" votes for easy and difficult respondents. They found no statistical difference between the proportions of "for" votes in these groupings.

However, neither of these two analyses mitigates the nonresponse bias in the data. With respect to the first analysis, none of the four demographic characteristics influenced how the survey respondents voted. Specifically, none of the statistical models (i.e., logit models) developed by Stratus reflect these four demographic features. Instead, the models reflect that income, recreation frequency, and several opinions about the alleged problem and the proposed solution influenced the voting. Stratus has not established that any of these primary influences on the voting patterns is correlated with these four demographic characteristics. Moreover, Stratus has not, and cannot demonstrate that the nonrespondents would have had patterns of use or income, or attitudes that match those of the survey respondents. Accordingly, the demographic analysis does little to mitigate nonresponse bias in these CV data.[53]

In terms of the second analysis, the underlying assumption is that the nonrespondents are similar to the late and/or difficult respondents. The presumption is that had the nonrespondents participated, they would be like the late and/or difficult respondents. Stratus is asserting that because the late and/or difficult to reach respondents do not have different voting patterns than do the early and/or easy

[53] OMB Circular A-4 reveals that "caution should be used in assessing the representativeness of the sample based solely on demographic profiles."

respondents, the overall vote would have not been any different. These assumptions and assertions do not address nonresponse bias. Clearly, there is something different about the nonrespondents relative to the late and/or difficult respondents. For reasons unknown, the nonrespondents chose to not participate in the study, despite several attempts to contact them. Given that the nonrespondents comprise well more than 40 percent of the sample households, their votes could have changed the outcome. Despite these limited analyses conducted by Stratus, nonresponse bias remains in the data.

The important implication of nonresponse bias in this assessment is the resulting inappropriateness of multiplying the average WTP from the CV survey by the total number of households in the 63 counties. Because the Stratus CV survey results reflect nonresponse bias, applying the WTP results to 1.4 million households is not appropriate, further underscoring the fact that the CV survey damage results are not reliable.

4.7 The damage estimates do not correspond to the proper economic baseline.

Finally, the Stratus damage estimate does not comport with the appropriate economic baseline conditions. In economic analysis, it is critical to establish the appropriate baseline conditions, which would be the aesthetic and ecosystem conditions but for the release of phosphorous from the application of poultry litter. Thus, it is necessary to net out the effects of other sources of phosphorous and their impact on water quality in the Illinois River System and Tenkiller Lake. The Stratus damage estimate does not reflect an appropriate baseline because it measures damages relative to hypothetical conditions in 1960. As Connolly, Sullivan, and Coale (2009) point out, there is no data to establish the conditions in 1960, and certainly no basis to argue that the photographs that were used to represent baseline actually reflected the conditions in 1960.[54] Furthermore, the damages estimated from the CV study reflect all of the past phosphorus in the Illinois River System and Tenkiller Lake.

[54] This means that rather than 48 years of the difference in aesthetics, the damages should be based on 27 years. Moreover, the past damages report prepared by Stratus confirms that the annual change in the visual aesthetics was constant over time. Thus, even if the damages were based on a valid estimate they are further inflated by this error.

According to the Stratus CV questionnaire, about 60 percent of the phosphorus is attributable to the poultry industry. However, as designed, the CV survey results reflect the purported value associated with ***all*** of the past phosphorus, not just the portion attributable to the poultry industry. Thus, the Stratus damages estimate, even if it were valid and reliable, does not correspond to the appropriate economic baseline conditions.

4.8 The Stratus CV survey does not conform to the NOAA panel guidelines.

The Stratus report contends that it has met the NOAA panel guidelines for conducting CV surveys. As our review has demonstrated, there are meaningful differences between the Stratus CV and the NOAA panel guidelines. In fact, the Stratus CV survey fails 16 out of the 24, or two-thirds, of the applicable guidelines. Moreover, the guidelines that they met are insufficient to overcome the serious flaws in the study. Table 4.10 below highlights these differences.

Table 4.10: Summary Table of NOAA Panel Guidelines

NOAA Panel Guideline	Sufficiently Addressed in Stratus CV Study?	Discussion of Items Not Sufficiently Addressed
Sample Size and Type	No	The different sample sizes for the base and scope versions influences the scope test results. With comparable sample sizes, the study would not meet the scope test guidelines.
Nonresponse Bias	No	The response rate is 52 percent, well below the guidelines set by NOAA and OMB. The nonresponse analysis does not address how the nonrespondents differ from the respondents in terms of the respondent opinions and experiences that influenced their votes on the program.
In-person Interviews	Yes	
Test for Interviewer Effects	No	Stratus conducted hotel pretests that purport to demonstrate that the in-person interviewer format did not affect voting patterns relative to a self-administered survey. However, the ballot box research described by Harrison (2007) indicates strong interviewer effects. Moreover, Stratus did not include design elements recommended by NOAA panel, such as a ballot box or mail-in survey component (Leggett, et al. 2003). In some in-person studies, individual interviewers can unduly sway the results (Leggett, et al. 2003).

NOAA Panel Guideline	Sufficiently Addressed in Stratus CV Study?	Discussion of Items Not Sufficiently Addressed
Data Reporting	Yes	
Careful Pretesting	No	The amount of pretesting does not correspond to careful pretesting. Careful pretesting would have documented the salient changes in the questionnaire over time, and the evolution of the bid levels used (Smith 2007). Moreover, the NOAA Panel guidelines indicate that careful pretesting will result in respondent comprehension and acceptance, which this study fails to demonstrate.
Conservative Design	No	The CV questionnaire is not balanced in terms of presenting information on the poultry industry and other sources of phosphorus. The CV questionnaire provides significant information dosing about the poultry impacts on water conditions before asking respondents' impressions of the resources, a sharp contrast to the earlier Stratus telephone survey.
WTP Elicitation Format	Yes	
Referendum Format*	Yes	
Accurate Description of Injury and Proposed Program	No	The information presented is at odds with available scientific information. Moreover, the damage estimate that results from the survey reflects injury from all past sources, not uniquely the poultry industry defendants. Because the described alum treatment does not distinguish the source of the phosphorus, the CV results are not relevant for damage assessment, as the NOAA Panel guidelines indicate.
Pretest Photographs	No	Although Stratus included photos in the pretests, they did not evaluate and report on any potential biases associated with various photographs. Moreover, the use of the photos exaggerates the purported injury because they do not portray the spatial and seasonal extent of algae conditions.
Reminder of Undamaged Substitutes**	Yes	
Adequate time lapse from incident	No	The NOAA Panel included this guideline to address frequent and biased media coverage of the environmental changes. The Attorney General filed this suit in 2005, and the media coverage has increased awareness of the algae conditions over the last year. In 2006, when Stratus conducted its telephone survey, less than 10 percent of the respondents volunteered any negative impressions of the resources. The majority of the 2006 respondents indicated that the river and lake were high-quality recreation resources, with clean and clear water.
Temporal Averaging	N/A	

NOAA Panel Guideline	Sufficiently Addressed in Stratus CV Study?	Discussion of Items Not Sufficiently Addressed
No answer option	No	The no-answer option was not included in this study. Harrison (2007) provides a dissection of the research on which Stratus relies when claiming that this guideline is irrelevant. Harrison demonstrates that the results are sensitive to interpretation and contrary to the information presented by Stratus.
Yes/No Follow-ups	No	Although the questionnaire included yes/no follow-ups to the WTP question, the results were not factored into the analysis, which was the intent of the NOAA Panel (Smith 2007). For example, about 40 percent of the respondents believed that the extra taxes would be used to clean up other lakes and rivers. Although these respondents are clearly thinking about a much broader suite of resources when they voted yes, and the analysis should have at least controlled for them.
Cross Tabulations	Yes	
Checks on Understanding and Acceptance	No	Although the questionnaire included questions that would reveal whether the respondents understood and accepted the scenario, the analysis of the results indicates that many respondents did not understand and/or accept the scenario.
Alternative Expenditure Possibilities	No	The "budget constraint," or reminder of alternative expenditure possibilities should be more than perfunctory, according to the NOAA Panel guidelines. The analysis conducted indicates that respondents did not consider their incomes during the hypothetical voting. With almost one third of respondents not paying income state income taxes, the budget constraint is ineffective.
Deflection of Transaction Value	No	The follow-up questions indicate that many respondents voted before the program because it would help the environment "in general." These respondents were not thinking of the specific resources at issue when they voted.
Steady State or Interim Losses	Yes	
Present Value of Interim Losses	Yes	
Advanced Approval	No	Stratus did not seek advance approval of the defendants.
Burden of Proof	No	The response rate does not meet the established thresholds. Many respondents did not understand or accept the scenario described by the interviewers. The scope test result depends entirely on the chosen statistical technique (see below).
Scope Test	No	The scope test result is a statistical artifact of the large sample size. Moreover, the scope test does not vary only one dimension, which leads to a confounding effect.

NOAA Panel Guideline	Sufficiently Addressed in Stratus CV Study?	Discussion of Items Not Sufficiently Addressed
* The hypothetical referendum format does not give the same results as a real referendum and provides no counter to hypothetical bias. ** Even though a substitute's reminder was included, more than 40 percent thought the hypothetical program would benefit other resources.		

Moreover, the following list details more differences between the Stratus CV and the NOAA panel guidelines:

- The survey design is not conservative as respondents were repeatedly dosed with information that either was factually incorrect, misleading, or unbalanced in its presentation. Respondents were given no information about potential economic tradeoffs, nor were respondents told about the potential uncertainty surrounding the proposed restoration project and the purported injuries to fish and other biota.
- The photographs are biased because they fail to remind respondents that the purported impacts would be seasonal and would only affect a portion of the lake. Plaintiffs do not make clear the portion of the lake that would be affected. The photographs also do not correspond to the stated conditions in the survey questionnaire.
- Budget constraint is merely perfunctory in the Stratus survey. Moreover, it is irrelevant for approximately one-third of the Stratus survey respondents who either got a full-refund or paid no state income taxes.

The Stratus survey does not fulfill the guidelines for a valid CV survey (nonresponse, scope, understanding, scenario acceptance). It does not include a valid scope test. Respondents routinely provided that answers that demonstrated that they ignored what was said in the survey interview as it related to the specific commodity that was to be valued. Moreover, the respondents' answers do not correspond to basic economic principles of the law of demand and income elasticity. Such failings are more than sufficient to indicate that the Stratus survey is not a valid basis for measuring damages.

4.9 References

51 *Fed. Reg.* 27751, 1986

Adamowicz, W., J. Louviere, and M. Williams. 1994. Combining Real and Stated Preference Methods for Valuing Environmental Amenities. *Journal of Environmental Economics and Management* 26: 271-292.

Arrow, Kenneth, Edward Leamer, Howard Schuman, and Robert Solow. 1994. ''Comments on NOAA Proposed Rule on Natural Resource Damage Assessment'', in *Comments on Proposed NOAA/DOI Regulations on Natural Resource Damage Assessment,* EPA, October 1994.

Arrow, K., R. Solow, P.R. Portnoy, E.E. Leamer, R. Radner, and H. Schuman. 1993. "Report of the NOAA Panel on Contingent Valuation." 58 *Fed. Reg.* 4601 *et. seq.* Jan. 15.

Bierman, Victor, Jr., Ph.D. 2009. "Expert Report of Victor J. Bierman, Jr. in the Matter of *State of Oklahoma, et al., Plaintiffs v. Tyson Foods, Inc., et al., Defendants.* United States District Court for the Northern District of Oklahoma. Case No. 05-CV-0329-GKF-SAJ. January 23.

Bishop Corr 0000126

Bishop undated. Notes on Hypothetical Bias. Bishop0001271_Notes on Hypothetical Bias.doc

Blumenschein, Karen, Magnus Johannesson, Glenn C. Blomquist, Bengt Liljas, and Richard M. O'Conor. 1998. "Experimental Results on Expressed Certainty and Hypothetical Bias in Contingent Valuation." *Southern Economic Journal* 65(1):169-177.

Carson, R.T. and T. Groves. 2007. "Incentive and informational properties of preference questions." *Environmental and Resource Economics.* 37:181–210.

Carson, R. T. 1997. "Contingent Valuation Surveys and Tests of Insensitivity to Scope", in R. J. Kopp, W. Pommerhene and N. Schwartz, eds., *Determining the Value of Non-Marketed Goods: Economic, Psychological, and Policy Relevant Aspects of Contingent Valuation Methods*. Boston: Kluwer, pp. 127–163.

Carson, R.T., Flores, N.E. and Meade, N.F. 2001. "Contingent Valuation: Controversies and Evidence." *Environmental and Resource Economics* 19: 173–210.

Carson, Richard T., Robert Cameron Mitchell, W. Michael Hanemann, Raymond J. Kopp, Stanley Presser, and Paul A. Rudd. 2003. *A Contingent Valuation Study of Lost Passive Use Values Resulting From the Exxon Valdez Oil Spill. Environmental and Resource Economics* 25: 257–286.

Carson, Richard T., Michael B. Conaway, W. Michael Hanemann, Jon A. Krosnick, Kerry M. Martin, Donald R. McCubbin, Robert Cameron Mitchell, and Stanley Presser. 1996. *The Value of Preventing Oil Spills to Natural Resources Along California's Central Coast.* Report prepared for the California Department of Justice, the Office of Oil Spill Prevention and Response in the California Department of Fish and Game, and the Damage Assessment Center of the National Oceanic and Atmospheric Administration. San Diego, CA: Natural Resource Damage Assessment, Inc.

Carson, Richard T., W. Michael Hanemann, Raymond J. Kopp, Jon A. Krosnick, Robert C. Mitchell, Stanley Presser, Paul A. Ruud, and V. Kerry Smith. 1994. *Prospective Interim Lost Use Value Due to DDT and PCB Contamination in the Southern California Bight.* Report prepared for the National Oceanic and Atmospheric Administration. La Jolla, CA: Natural Resource Damage Assessment, Inc. September 30.

Champ, Patricia A., and Richard C. Bishop. 2001. "Donation Payment Mechanisms and Contingent Valuation: An Empirical Study of Hypothetical Bias." *Environmental and Resource Economics* 19:383-402.

Champ, Patricia A., Richard C. Bishop, Thomas C. Brown, and Daniel W. McCollum. 1997. "Using Donation Mechanisms to Value Nonuse Benefits from Public Goods." *Journal of Environmental Economics and Management* 33:151-162.

Code of Federal Regulations. 43 CFR Part 11.

Connolly, John P. 2009. "Expert Report of John P. Connolly: Illinois River Watershed Water Quality and Source Assessment." Prepared for: Illinois River Watershed Joint Defense Group. Quantitative Environmental Analysis, LLC: Montvale, NJ. January 30.

Connolly, John P., Ph.D., Timothy Sullivan, Ph.D., and Frank Coale, Ph.D. 2009. Expert report in the matter of *State of Oklahoma v. Tyson Foods, et al.* Evaluation of Hypothetical Remediation Strategy Presented in Stratus Contingent Value Study Illinois River Watershed. March.

Cooke, Dennis. 2008. Expert testimony of Dr. Dennis King in the Matter of *W.A. Drew Edmondson, in his capacity as Attorney General of the State of Oklahoma and Oklahoma Secretary of the Environment C. Miles Tolbert in his capacity as the Trustee for Natural Resources for the State of Oklahoma vs. Tyson Foods, et al.* January 30. In the United States District Court for the Northern District Of Oklahoma. CV-00329-TCK-SAJ

Desvousges, W.H., D.J. MacNair, and G.A. Smith. 2000. *Lower Fox River and Bay of Green Bay: Assessment of Potential Recreational Fishing Losses and Restoration Offsets*. Durham, NC: Triangle Economic Research.

Desvousges, W.H., S.P. Hudson, and M.C. Ruby. 1996. "Evaluating CV Performance: Separating the Light From the Heat." In *The Contingent Valuation of*

Environmental Resources: Methodological Issues and Research Needs, eds. D.J. Bjornstad and J.R. Kahn. Brookfield, VT: Edward Elgar Publishing Limited.

Desvousges, William H., F. Reed Johnson, Richard W. Dunford, Kevin J. Boyle, Sara Hudson, and K. Nicole Wilson. 1993. "Measuring Natural Resource Damages with Contingent Valuation: Tests of Validity and Reliability." In *Contingent Valuation: A Critical Assessment*, edited by Jerry A. Hausman. Amsterdam: North Holland Publishers.

Diamond, P.A. and J.A. Hausman. 1994. "Contingent valuation: is some number better than no number?" *Journal of Economic Perspectives* 8(4): 45–64.

Diamond, P.A. and J.A. Hausman, ed. 1993. "On Contingent Valuation Measurement of Nonuse Values." In *Contingent Valuation: A Critical Assessment.* Amsterdam: Elsevier Science Publishers B.V.

Dunford, Richard, Alicia R. Gable, Sara P. Hudson, William H. Desvousges, and F. Reed Johnson. 1996. *A Review of the 1996 Contingent Valuation Study on the Value of Preventing Oil Spills Along California's Central Coast.* Durham, NC: Triangle Economic Research. October.

Ethier, R.G., Poe, G.L., Schulze, W.D., Clark, J., 2000. "A comparison of hypothetical phone and mail contingent valuation responses for green-pricing electricity programs." *Land Economics* 76: 54–67.

Fisher, A.C. 1996. "The conceptual underpinnings of the contingent valuation method." In *The Contingent Valuation of Environmental Resources: Methodological Issues and Research Needs*, eds. D.J. Bjornstad and J.R. Kahn. Brookfield, VT: Edward Elgar Publishing Limited.

Freeman III, A.M. 1993. *The Measurement of Environmental and Resource Values.* Washington, DC: Resources for the Future.

Green, D., K.E. Jacowitz, D. Kahneman, and D. McFadden. 1998. "Referendum contingent valuation, anchoring, and willingness to pay for public goods." *Resource and Energy Economics* 20(2): 85–116.

Harrison, Glenn W. 2007. "Making Choice Studies Incentive Compatible." In *Valuing Environmental Amenities Using Stated Choice Studies: A Common Sense Approach to Theory and Practice*, ed. Barbara Kanninen. Dordecht, The Netherlands: Springer.

Horowitz, John. 2000. "Discussion of Theory and Design of Stated Preference Methods." Stated Preference: What Do We Know? Where Do We Go? A Workshop Sponsored by the USEPA's National Center for Environmental Economics and National Center for Environmental Research. Washington, DC: October 12.

Kahneman, Daniel and Jack Knetsch. 1992. "Valuing Public Goods: The Purchase of Moral Satisfaction." *Journal of Environmental Economics and Management* . 22: 57-70.

Kanninen, B.J., 1995. "Bias in discrete response contingent valuation." *Journal of Environmental Economics and Management* 28: 114–125.

Kemp, M.A. and C. Maxwell. 1993. "Exploring a Budget Context for Contingent Valuation Studies." *Contingent Valuation: A Critical Assessment*, ed. J.A. Hausman. Amsterdam: Elsevier Science Publishers B.V.

King, Todd. 2009. Expert testimony of Todd King in the Matter of *W.A. Drew Edmondson, in his capacity as Attorney General of the State of Oklahoma and Oklahoma Secretary of the Environment C. Miles Tolbert in his capacity as the Trustee for Natural Resources for the State of Oklahoma vs. Tyson Foods, et al.* January 30. In the United States District Court for the Northern District Of Oklahoma. CV-00329-TCK-SAJ

King, Todd. 2008. "Identification and Evaluation of Viable Remediation Alternatives to address Injuries related to Land Disposal of Poultry Waste within the Illinois River Watershed." May 15.

Leggett, Christopher G., Naomi S. Kleckner, Kevin J. Boyle, John W. Duffield, and Robert Cameron Mitchell. 2003. "Social Desirability Bias in Contingent Valuation Surveys Administered Through In-Person Interviews." *Land Economics* 79(4): 561-575.

Mathews, K.E., M.L. Freeman, and W.H. Desvousges. 2007. "How and How Much? The Role of Information in CE Questionnaires." In *Using Choice Experiments to Value Environmental Amenities*, Barbara Kanninen, ed. Boston: Kluwer Academic Publishers.

Mitchell, Robert Cameron and Richard T. Carson. 1989. *Using Surveys to Value Public Goods: The Contingent Valuation Method.* Washington, DC: Resources for the Future, Inc.

Morey, Edward 2006. "Some Initial Thoughts on the Phone Survey." Draft. November 28.

OMB 2006. *Standards and Guidelines for Statistical Surveys.* Office of Management and Budget. Washington, DC. September.

Poe, Gregory L., Jeremy E. Clark, Daniel Rondeau, and William D. Schulze. 2002. "Provision Point Mechanisms and Field Validity Tests of Contingent Valuation." *Environmental and Resource Economics* 23:105–131.

Presser, S., J. Blair, and T. Triplett. 1992. "Survey Sponsorship, Response Rates, and Response Effects." *Social Science Quarterly* 73(3): 699–702.

Rea, Louis M. and Richard A. Parker. 2005. Designing and Conducting Survey Research: A Comprehensive Guide. San Francisco: Josey-Bass Publishers.

Rossi, P., J. Wright, and A. Anderson. 1983. *Handbook of Survey Research*. New York: Academic Press, page 215.

Rowe, R.D., R. d'Arge, and D.S. Brookshire. 1980. "An Experiment on the Economic Value of Visibility." *Journal of Environmental Economics and Management* 7:1–19.

Schläpfer, F. 2008. "Contingent valuation: A new perspective." *Ecological Economics* 64(4): 729-740.

Schläpfer, F. and I. Bräuer. 2007. "Theoretical incentive properties of contingent valuation questions: do they matter in the field." *Ecological Economics* 62: 451–460.

Schulze, W.D., G.H. McClelland, J.K. Lazo, and R.D. Rowe. 1998. "Embedding and Calibration in Measuring Non-Use Values." *Resource and Energy Economics* 20, 163–178.

Schuman, H. and S. Presser. 1981. *Questions and Answers in Attitude Surveys*. New York: Academic Press.

Slovic, P., B. Fischoff, and S. Lichtenstein. 1982. "Facts versus Fears: Understanding Perceived Risks." In *Judgment Under Uncertainty: Heuristics and Biases*, eds. D. Kahneman, P. Slovic, and A. Tversky. Cambridge: Cambridge University Press.

Smith, V. Kerry. 2007. "Judging Quality." In *Valuing Environmental Amenities Using Stated Choice Studies: A Common Sense Approach to Theory and Practice*, ed. Barbara Kanninen. Dordecht, The Netherlands: Springer.

Smith, V.K. 1986. "To Keep or Toss the Contingent Valuation Method." In *Valuing Environmental Goods: An Assessment of the Contingent Valuation Method*, eds. R.G. Cummings, D. Brookshire, and W.D. Schulze. Totowa, NJ: Rowman and Allanheld.

Smith, V. K. and L. Osborne. 1996. "Do Contingent Valuation Estimates Pass a Scope Test? A Meta Analysis." *Journal of Environmental Economics and Management* 31: 287–301.

Stratus 2007. "Oklahoma Watershed Short Telephone Survey Report (Draft)." March 29. (Stratus0027772_STS.Draft_NEW.doc)

Swait, J., J.J. Louviere, and M. Williams. 1994. "A Sequential Approach to Exploiting the Combined Strengths of SP and RP Data: Application to Freight Shipper Choice." *Transportation* 21:135–152.

Tourangeau, Roger, Lance J. Rips, and Kenneth Rasinski. 2000. *The Psychology of Survey Response.* Cambridge: Cambridge University Press.

Tversky, A. and D. Kahneman. 1981. "The Framing of Decisions and the Psychology of Choice." *Science* 211(January 30):453–458.

Vossler, Christian A. and Joe Kerkvliet. 2003. "A Criterion Validity Test of the Contingent Valuation Method: Comparing Hypothetical and Actual Voting Behavior for a Public Referendum." *Journal of Environmental Economics and Management* 45:631–649.

Vossler, Christian A., Joe Kerkvliet, Stephen Polasky, and Olesya Gainutdinova. 2003. Externally Validating Contingent Valuation: An Open-Space Survey and Referendum in Corvallis, Oregon. *Journal of Economic Behavior and Organization* 51:261–277.

Weinstein, M.C. and R.J. Quinn. 1983. "Psychological Considerations in Valuing Health Risk Reductions." *Natural Resources Journal* 23(3):659–673.

Whittington, D., G. Cassidy, D. Amaral, E. McClleland, H. Wang, and C. Poulous. 1994. "The Economic Value of Improving the Environmental Quality of Galveston Bay." The Galveston Bay Estuary Program.

5. BIASES RESULTING FROM STATISTICAL AND ECONOMETRIC ANALYSES

In this section, we examine the statistical and econometric analyses presented in the Stratus CV Report. We isolate a number of important biases in the reported scope test estimates of WTP, as well as a number of inherent violations of fundamental principles. Our focus is on the robustness, or lack thereof, of the reported estimates. None of the WTP estimates generated as part of this robustness analysis should be viewed as a basis for an alternative measure of damages.

5.1 Stratus employs a non-parametric estimator of WTP, resulting in unreliable WTP estimates.

The Stratus Report overestimates the WTP that can properly be based upon the survey data. The WTP measures presented in the Stratus report use the nonparametric ABERS estimator. We tested the robustness of the ABERS estimator by employing, instead, the nonparametric Turnbull estimator and found that the Turnbull estimator produced more conservative estimates of WTP. Although they are more conservative, as we show below, neither of these approaches produces valid a WTP estimate.[55]

The ABERS and Turnbull estimators assume that the probability that WTP is below a certain dollar amount increases as that dollar amount grows. For example, if there is a 50-percent likelihood that respondents' WTP will be $10 for a specified restoration effort, that probability would be expected to decline (hypothetically to perhaps 20%) as the price tag increases to $25. This relationship is known as a monotonically increasing cumulative distribution function for WTP. Both estimations rely on a recursive process, comparing frequencies of "no" votes for two bid amounts at

[55] These nonparametric approaches are more reliable than parametric estimators because they avoid assumptions regarding the distribution of WTP between bid amounts. Parametric estimators interpolate data between bid amounts so that every dollar amount is associated with some number of people who hypothetically exhibit that WTP. For example, although we do not have any vote data for bids of $172.50, parametric estimators assume that some frequency of people who said "No" to a bid of "$205" would say "Yes" to $172.50. This interpolated frequency is extremely sensitive to the distributional assumptions made and results in unreliable WTP estimates.

a time.[56] If the higher bid amount is always associated with a higher frequency of "no" votes, the ABERS and Turnbull estimators yield the same WTP.

When this is not the case, the ABERS and Turnbull estimators proceed by taking the weighted average of frequencies for the two bids. However, the ABERS procedure assigns this new frequency to both bids, while the Turnbull estimator assigns this new frequency to the lower bid and effectively drops the higher bid. By continuing to weight the higher bid, the ABERS estimator artificially creates a lower bound WTP for people who reject the next higher bid.

For example, suppose the "Yes" vote frequencies for $80 and $125 contradict the existence of a monotonically increasing cumulative distribution function. If the new ABERS frequency of a "Yes" vote for both these bids is 0.50 and the empirical frequency of a "Yes" vote for a $205 bid is 0.30, the ABERS estimation procedure assumes that the difference of 0.20 is caused by people having valuations between $125 and $205. In contrast, the Turnbull estimator drops the $125 bid entirely so that the $125 does not act as a "bottom floor" for WTP. In other words, the ABERS estimator places artificial lower bounds at arbitrary bid amounts, resulting in WTP estimates that are biased upward.

In describing its WTP estimation procedure, the Stratus report states: "… the estimated mean converges to the true mean of the distribution from below, meaning that the estimated mean underestimates the true mean in finite data sets."[57] This statement wrongly implies that the ABERS WTP always underestimates the hypothetical "true" WTP, which is achieved as sample sizes approach infinity. This is only accurate when the original method of smoothing the cumulative distribution function is correct. A more accurate description is that the ABERS estimator will arrive at a larger WTP in response to increases in the maximum bid, as is true with the Turnbull estimator. However, the WTP derived from the ABERS estimator will

[56] The cumulative distribution function is the probability that WTP is lower than a given bid, or that a respondent will cast a "No" vote. The distribution of "Yes" Votes for a Population is thus one minus the cumulative distribution function at every point, so that an increasing cumulative distribution function is equivalent to a decreasing distribution of "Yes" votes.

[57] The "mean" refers to WTP, which is calculated as the mean of the cumulative distribution function, Stratus Report, Vol I, p. 166.

consistently be equal to or higher than the WTP derived from the Turnbull estimator for any finite sample. There is no statistical justification for the implication in Stratus' report that the ABERS estimator would systematically underestimate true WTP. If the ABERS estimate of the cumulative distribution overemphasizes larger WTP by creating "bottom floors," the ABERS estimates and resulting WTP calculations will also result in overestimates, as is the case here.

There are also differences in the calculation of standard errors to form confidence intervals. Following the Stratus report, we use a jackknife bootstrap[58] to obtain standard errors for ABERS estimates, as well as confidence intervals for empirical cumulative distributions throughout this analysis. However, since the jackknife procedure requires more than one primary sampling unit in a stratum, the structure of the survey makes it impossible to use the jackknife procedure on many subpopulations of interest, such as passive versus active users, in the survey. When this is the case for the ABERS estimate, we leave the appropriate column/row blank. For the Turnbull estimator, we use asymptotic theory throughout to generate standard errors for every subpopulation.[59]

In light of the strong biases present throughout the contingent valuation method employed by Stratus that generate higher WTP estimates, it is especially important that the valuation methods employed avoid contributing further positive bias. We examined the ABERS and Turnbull WTP, using 95% confidence intervals for the entire population and the empirical distribution of "Yes" votes for that population expressed as "Pr" or the "probability" of a Yes vote at the associated bid level.[60] The distribution of "Yes" votes (equivalent to 1 minus the cumulative distribution function) is not monotonically decreasing. As a result, WTP derived using the ABERs estimator adopted by Stratus is greater than the estimated WTP using the Turnbull estimator, as indicated in Table 5.1.

[58] A jackknife procedure estimates standard errors of estimates by repeatedly calculating the estimator, leaving out a single different observation each time. When there is a single primary sampling unit in a survey stratum, the jackknife has no way to re-estimate and thus cannot be implemented.

[59] This is the most common procedure for estimating standard errors. For specific use in non-market valuation see Haab T.C. and McConnell K.E. 2002. "Valuing Environmental and Natural Resources."

[60] The cumulative distribution function is the probability that WTP is lower than a given bid, or that a respondent will give a "No" vote. The distribution of "Yes" Votes for a Population is thus one minus the cumulative distribution function at every point so that an increasing cumulative distribution function is equivalent to a decreasing distribution of "Yes" votes.

While the ABERS estimate of WTP is $184.44 (based on a 95% confidence interval of $165.72 to $203.38), the Turnbull estimate is almost $6 lower, at $176.78 (based on a 95% confidence interval of $160.09 to $193.42). As previously explained, these divergent WTP estimates are due to the difference in smoothing procedures for the ABERS and Turnbull estimators. Any WTP estimates for subpopulations of respondents who display a similarly inconsistent decline in "yes" votes as the bid increases (e.g., an *increase* in the number of respondents voting "yes" when the bid increases from $80 to $125 as displayed for the total population in Table 5.1) will similarly result in overestimates of WTP using ABERS, as discussed in the next section of this report.

Table 5.1: Application of ABERS and Turnbull Estimators Producing Different WTP from Same Data

Empirical Distribution Of "Yes" Votes for Population			
	Bid Amount	**Pr (Yes|Bid)**	
	$10.00	0.815	
	$45.00	0.701	
	$80.00	0.602	
	$125.00	0.615	
	$205.00	0.435	
	$405.00	0.342	
	WTP	**95% lower**	**95% upper**
ABERS	$184.55	$165.72	$203.38
Turnbull	$176.78	$160.09	$193.42

5.2 Analysis of Subgroups of Respondents

In addition to examining Stratus' results for the entire surveyed population, we compared estimated WTP for several key subpopulations to determine whether the results are consistent with economic logic. This is a standard approach to evaluating the robustness of claimed survey results. The subpopulations tested were based upon: 1) the nature of the respondent's use of the natural resources (either active or passive), 2) respondent's perception of the effectiveness of the proposed alum treatment, 3) difficulty respondent would likely have in paying the proposed alum tax, and 4)

respondent's opinions regarding levels of state spending on pollution. Comparison of active and passive users suggests that the survey is flawed because its results run contrary to fundamental economic logic: using the Turnbull estimator, passive users of Tenkiller Lake have a higher WTP for its restoration than do active users. The other sub-groups of respondents demonstrate such marked differences in WTP that the survey cannot be used to represent the views of the population at large without first knowing how that general population is distributed by each defining characteristic.

Passive Versus Active Users

Basic economic principles of demand dictate that as price increases, consumer demand will decrease. In economics, this principle is known as the law of demand. However, particular characteristics of the "buyers" should be expected to influence the value they place on the purchased good, and thus their WTP for it. These characteristics include both the money available to the buyer to make the purchase and the use which the buyer will make of the purchased good.

Extensive academic literature has employed or examined survey methods to estimate the WTP for various types of environmental quality improvements. Kriström and Riera (1996) and Hokby and Soderqvist (2003) review several European studies and find that the income elasticity of WTP for environmental quality improvements is almost always positive, but on average is less than one. These results are consistent with what economists would describe as "normal goods," which are not viewed as luxuries. This conclusion has been confirmed by Henderson (2008) and Fisher and Waschik (2002). Because environmental quality is a "normal good", as household income rises, the WTP for improvements to environmental quality also rises (Organization for Economic Co-operation and Development 2001).

The WTP for active users of Tenkiller Lake contradicts this intuition and economic logic. Question 14 asks whether respondents have ever visited the Illinois River and Question 15 elicits the same information for Tenkiller Lake. Respondents who have visited either the river or the lake are considered "Active" users of that area. Among active users of Tenkiller Lake, a higher portion of Stratus respondents were willing to pay $405 than were willing to pay $205 for the restoration program.

We examined the distribution of "Yes" votes, conditional on both bid amount and WTP for passive and active users of both the river and the lake. As expected, active users of the Illinois River have a higher WTP than passive users of that same resource. These results appear in Appendix C. In contrast, active users of the lake have a ***lower*** Turnbull WTP ($135) than do passive users of that same resource ($142). This is due to the non-monotonically decreasing empirical distribution for active users of the lake. In particular, the distribution increases from 0.46 to 0.49 for bids $205 and $405, respectively. As previously explained, while the ABERS estimator takes a weighted average of these proportions and assigns them to both bids, the Turnbull estimator assigns this weighted average to the $205 bid and ignores the $405 bid. This type of behavior in the data illustrates why the profession generally prefers the Turnbull estimator to the ABERS estimator, especially in light of the many upward biases in the Stratus survey. The large difference between bids in this case is the driving factor for the large difference in WTP estimates and points to the unreliability of the valuation technique.

Table 5.2: Comparison of WTP for Active v. Passive Users of Tenkiller Lake

Q15: Have you ever visited Tenkiller Lake?

	No (Passive Users)			Yes (Active Users)		
	WTP	**Lower 95%**	**Upper 95%**	**WTP**	**Lower 95%**	**Upper 95%**
ABERS	$149.89			$216.83		
TRNBL	$142.08	$114.13	$170.04	$135.00	$126.28	$143.72
	Pr(Yes|Bid)	Lower 95%	Upper 95%	Pr(Yes|Bid)	Lower 95%	Upper 95%
$10	0.76	0.76	0.76	0.87	0.87	0.87
$45	0.69	0.69	0.69	0.72	0.72	0.72
$80	0.55	0.55	0.55	0.65	0.65	0.65
$125	0.60	0.60	0.60	0.64	0.64	0.64
$205	0.40	0.40	0.40	0.46	0.46	0.46
$405	0.20	0.20	0.20	0.49	0.49	0.49

Effectiveness of Alum Treatment

The second subpopulation group tested consisted of respondents who believe that the alum treatment will be "Not", "Slightly", or "Moderately" effective as contrasted with those who think the alum treatment will be "Very" or "Extremely" effective. Not surprisingly, respondents viewing the alum treatment as "Very" or "Extremely" effective

had a WTP more than two times that of the other respondents. More than 60 percent of respondents concluded (after the considerable dosing described above) that the alum treatment would be "Very" or "Extremely" effective; they were willing to pay approximately $235.29 using the Turnbull estimate. In contrast, 35.5% of respondents believed that the alum treatment would be "Not", "Slightly", or "Moderately" effective; these respondents were only willing to pay $100.88. These differences are statistically significant. Detailed results appear in Appendix C, but it is clear that WTP for the restoration program depends heavily upon what scientific evidence is provided and the conclusions that people draw from it. Under these circumstances, the bias introduced through Stratus' survey presentation makes it impossible to draw any reliable conclusions.

State Spending on Pollution

The third subpopulation group we examined separated respondents who thought the state should spend "Less" or the "Same" on pollution from those who thought it should spend more. The former group represents 31.58% of the population; while 66.29% believe that the state should spend more to control pollution. Not surprisingly, those respondents who think the state should not increase its spending had a lower WTP ($119.96 using Turnbull) than did the overall population of respondents ($176.78 using Turnbull or $184.55 using ABERS estimators). This difference, for which the details appear in Appendix C, is statistically significant. Because 52.44 percent of all respondents either believe that the state should not increase spending on pollution or believe alum treatment will be no more than moderately effective, or believe both of these things, the views of the general population must at least roughly correspond with these distributions for the survey results to be meaningful in predicting the broader response. Stratus provides no method to extrapolate from these individual characteristics of the sample population to the larger population whose WTP is being estimated.

5.3 Implied bid and income elasticities are inconsistent with economic theory.

Two standard economic measures known as "elasticities" may also be computed and evaluated for compliance with established economic principles. The elasticity of demand is a well-established economic principle that measures the responsiveness of change in demand for a good or service relative to a change in its price. Income elasticity measures responsiveness to changes in the level of purchasers' incomes.

The elasticity measures how responsive demand is to a price increase or decrease. When the change in demand corresponds to a change in price, demand for the good is said to have "unitary elasticity." If, for example, the price increases 10 percent, the demand for the good will fall 10 percent. If the change in demand is greater than the change in price, demand is said to be "elastic." This occurs when the demand changes more than the price does. For example, a 20 percent decrease in demand that follows a 10 percent increase in price reflects elasticity. By contrast, a change in demand that is smaller than a corresponding price change is said to reflect "inelastic" demand. Certain necessities, such as gasoline and household water, cannot be easily be replaced or dispensed with and are less responsive to price changes than are discretionary goods (Scheierling, Loomis, and Young 2006). These necessities exhibit inelastic demand; consumers must buy them regardless of change in price.

As previously described, a price increase ordinarily results in a decrease in the quantity of an ordinary good purchased by consumers. This relationship is illustrated with the commonly-accepted downward sloping demand curve. A wide body of contingent valuation studies observe that income is positively associated with WTP for aesthetic public goods.[61] In a meta-analysis based on 46 contingent evaluation studies across six continents, Jacobsen and Hanley (2008) conclude that there is a significant positive effect of both personal income and national GDP on WTP for species and

[61] Bateman, I.J., and Langford, I.H. 1997. "Non-users' Willingness to Pay for a National Park: An Application and Critique of the Contingent Valuation Method." *Regional Studies* 31(6): 571-582; Ph. Le Goffe. 1995. "The Benefits of Improvements in Coastal Water Quality: A Contingent Approach." *Journal of Environmental Management* 45: 305-317; Alberini, A., Rosato, P., Longo, A., Zanatta, V. *Information and Willingness to Pay in a Contingent Valuation Study: The Value of S. Erasmo in the Lagoon of Venice."* Nota Di Lavoro 19.2004. February 2004.

habitat conservation.[62] Where a significant impact of income on WTP is not exhibited, researchers suspect that the reason is flaw in the survey design.[63]

In this case, it is possible to compute from the Stratus CV results: (1) the elasticity of demand with respect to changes in the hypothetical prices—the cost of the alum treatment program—and (2) the elasticity of demand with respect to changes in respondent income. The bid design, which randomly assigned different costs of the alum treatment program among respondents, and the inclusion of respondent income information in the Stratus CV survey, enable the calculation of these two important elasticities. Thus, an evaluation similar to the scope test determines the sensitivity of the "votes" to changes in the bid price. Because the item being purchased is a hypothetical water quality improvement rather than a unit of goods, the elasticity is measured with quantity represented by the probability of voting yes on the restoration program.[64] In the Stratus CV survey, the good being purchased is environmental quality, which is an ordinary (rather than luxury) good and should exhibit normal elasticity.

[62] Jacobsen, B.R., and Hanley, N. 2008. "Are there income effects on global willingness to pay for biodiversity conservation?" *Environmental and Resource Economics (*August).

[63] It has been argued that the income elasticities in CV surveys are too low to accord with economic intuition. In a meta-analysis of 64 studies reporting on 83 different valuation scenarios involving environment-related public goods, Schläpfer (2005) found that the low income effects may be an artifact of the survey method. Schläpfer, F. 2006. "Survey protocol and income effects in the contingent valuation of public goods: A meta-analysis." *Ecological Economics* 57: 415-429.

[64] This technique has been repeatedly recognized in the literature. Alberini, A., Kanninen, B., and Carson, R.T. 1997. "Modeling Response Incentive Effects in Dichotomous Choice Contingent Valuation Data." *Land Economics* 73(3): 309-24. Eckerlund, I., Johannesson, M., Johansson, Per-Olov., Tambour, M., Zethraeus, N. 1995. "Value for money? A contingent valuation study of the optimal size of the Swedish health care budget." *Health Policy* 34: 135-143.

Table 5.3 contains the elasticity calculations for both the base and scope versions of the Stratus survey. As the bid amounts increase, the percentage change in the bid amount is calculated. Similarly, the corresponding change in the percentage of respondents who vote in favor of the program (the quantity) is calculated. The comparison of these percent changes reflects the elasticity. When the bid goes up from $10 to $45, this represents a 350 percent price increase. However, in response to this price increase, the quantity demanded (represented by the probability of a "Yes" vote as reported by Stratus) falls only 14 percent, reflecting inelastic demand. Because all of the elasticity calculations are less than 1, the respondents' demand for improvements to the river and lake are inelastic. This result is inconsistent with expectations based upon the extensive literature previously described. Even more troubling, in the base survey result, the demand (probability of voting "Yes") actually increases rather than declining in response to an increase in price (the Bid price). When the bid rises from $80 to $125 (a 60 percent price increase) the proportion of Respondents voting "Yes" to the expenditure simultaneously increases. Such results violate accepted economic tenets.

Table 5.3: Base Questionnaire Bid Elasticity

Bid	Pr(Yes) (%)	% Change Quantity	% Change Bid	Bid Elasticity
$10	81.5			
$45	70.1	-14.0	350.0	-0.040
$80	60.2	-14.1	77.8	-0.182
$125	61.5	2.2	56.3	0.038
$205	43.5	-29.3	64.0	-0.457
$405	34.2	-21.4	97.6	-0.219

To more precisely examine the relationship between respondents' income and their WTP, we divide the income distribution into quartiles and separately estimate WTP for each quartile.[65] As displayed in Table 5.4, there is no consistent positive relationship between income and WTP as theory would predict, regardless of which estimator is employed.

[65] Those respondents with coded incomes at or over $ 99,999,999,998 are dropped from this analysis.

Table 5.4: WTP and Confidence Intervals for Income Quartiles

	ABERS	Turnbull	Lower 95%	Upper 95%
$60,000<y<$600,001	$186.11	$173.36	$97.73	$248.99
$33,000<y<$60,000	$181.44	$175.68	$110.17	$241.20
$18,000<y<$33,000	$186.94	$183.80	$151.82	$215.77
$0<y<$18,000	$187.14	$166.10	$131.75	$200.45

The wide confidence intervals for WTP within each income quartile indicate that there is large variation in choice about the restoration project based upon respondent income. These differences in WTP by income quartile are not statistically significant. The results persists regardless of how the income distribution is disaggregated (quartiles, quintiles or sextiles), and thus is not an artifact of arbitrary income grouping. Division into quintiles does not show a consistently positive relationship between income and WTP, regardless of which estimator is used. When respondent income is divided into sextiles, those in the lowest sextile (income of less than $13,000 per year) have the highest WTP, while those in the next sextile (income of $13,000 to $23,000 per year) have the lowest WTP. There is no sound theoretical basis for explaining why people would exhibit increasing tolerance to pollution as they grow wealthier and, yet, that is what the Stratus results would suggest as reported in Appendix C.

We also examined the relationship between WTP and income by estimating logit models.[66] In this model, the dependent variable is a binary representation of a vote, "1" being "Yes" and "0" being "No." Following the Stratus report, we use a

[66] *In all cases, we employed Stratus' logit specification to predict respondents' votes controlling for: (1) Bid amount, (2) Log income, (3) an indicator for whether you visited Tenkiller Lake or Illinois River more than six times in the last year, (4) how important the respondent thought it was to reduce state income taxes, (5) whether the state should spend less money on pollution, (6) whether the respondent lives a greater distance than the median respondent distance, (7) how serious the respondent considers the problem to be, (8) whether the respondent thinks the alum treatment would take place without the ban, (9/10) whether the respondent thinks that natural recovery will be slower or faster than stated, (11) whether the respondent believes that alum treatment will be effective, (12) whether the respondent believed the tax would be used to clean other rivers or lakes, (13) respondents' trust in scientists and the Oklahoma government, (14) preferred method for funding environmental programs, (15) whether respondents considered themselves "environmentalists", (16) whether respondents felt pushed to vote for the program, (17) whether respondents paid OK state taxes, (18) whether respondents believed actual program cost would exceed the stated cost.

jackknife bootstrap to obtain standard errors for logit coefficient estimates throughout this analysis.[67]

Using the full logit model as specified in the Stratus Report, we estimate the income elasticity of "Yes" vote proportions to be 0.120 with a standard error of 0.03, evaluated at the mean of all independent variables. The positive elasticity indicates that as income increases the probability of a "yes" vote also increases. However, when we disaggregate income further, we see that the income elasticity is highly unstable. In fact, for certain income groupings income elasticity is once again *negative,* indicating that as income increases the probability of a "yes" vote for the proposed restoration program actually *decreases.* Table 5.5 reports the initial income groupings we created from wealthiest to poorest, the mean income for each of those groupings, and the estimated income elasticity of "Yes" votes at that point. Because these results are consistent for income quartiles, quintiles and sextiles, they clearly are not an artifact of arbitrary grouping.

Table 5.5: Estimated Income Elasticities by Income Groupings Using Logit Model

Grouping	Mean Income	Elasticity
Quartile	$330,000.50	0.2002
	$46,500.00	0.177
	$25,500.00	-0.397
	$9,000.00	0.158
Quintile	$332,000.50	-0.0399
	$53,500.00	-0.446
	$35,000.00	0.745571
	$21,000.00	-0.20074
	$7,500.00	0.224
Sextile	$335,000.50	-0.09302
	$60,000.00	0.572
	$41,500.00	1.51442
	$28,000.00	-0.58784
	$18,000.00	-0.6964
	$6,500.00	0.160972

[67] The results reported here are equivalent to those obtained from probit specification.

These results raise serious questions about the validity of the Stratus CV study. The basic relationship between income and WTP is an established tenet of economics, which is violated by the Stratus reported results.

5.4 Recoding of Base Survey Data

Section 4 of this report highlights many instances in which respondents did not understand the CV scenario, did not accept the "facts" presented by the interviewers, were not certain of their vote, or demonstrated inconsistencies in logic within their responses. This section describes additional analyses of the data when re-coded to address these issues. The following recoding procedures were applied:

- If respondents voted "Yes", but also indicated that they believed the new tax revenue would be used to pay for alum treatments to clean up other rivers and lakes in Oklahoma (contrary to the survey instructions), their votes were recoded as "No". ["Pay Other River/Lake Recode"]
- If respondents voted "Yes", but also thought the alum treatments might occur without ban, their votes were recoded as "No". ["Alum Without Ban Recode"]
- If respondents voted "Yes" but were unsure of their votes (moderately sure or less), their votes were recoded as "No". ["Certainty Recode"]
- If respondents voted "Yes", but paid no state income tax or received a full refund and thus would bear no cost for the program, their votes were recoded as "No". ["Income Tax Recode"]
- If the respondents voted "Yes", but thought restoration would be faster than described for the lake or the river, their votes were recoded as "No". ["Faster – River Recode and Faster – Lake Recode"]

The number and percent of "Yes" votes that were recoded for each of these reasons is presented in Table 5.6. The number of recodes for the base survey is presented relative to the Stratus' original dataset. The Alum Without Ban recode resulted in the smallest number of recodes, changing only 20% of the originally coded "yes" votes to "no." Cumulatively, 519 "Yes" votes (80%) are changed to "No" votes when employing all six recodes.

Table 5.6: Recode of "Yes" Vote For Inconsistency in Logic

Dataset	# "Yes" Votes	# Recodes	% Recoded
Stratus Original	**647**	**0**	**0.00%**
Recode: Alum Without Ban	517	130	20.09%
Recode: Certainty	485	162	25.04%
Recode: Faster - River	505	142	21.95%
Recode: Faster - Lake	504	143	22.10%
Recode: Pay Other River/Lake	358	289	44.67%
Recode: Income Tax	389	258	39.88%
Combined Recodes	128	519	80.22%

5.5 WTP Calculated with Recoded Data

Estimates of WTP change dramatically when the recoded data are employed. Stratus' WTP developed using the ABERS estimator was $184.55, which was corrected to $176.78 using the Turnbull estimator. These differences in WTP, which are statistically significant on their own, are highly sensitive to each of the data recodes described above. The results for each of the recodes are set forth in Table 5.7. Individual adjustments result in anywhere from a 15% decline in WTP estimates (Alum Without Ban) to a 44% decline (Other River/Lake). If even one of these recodes is employed, Stratus' damage estimate would correspondingly decline significantly.

Table 5.7: Turnbull WTP Estimates for Original and Recoded Data[68]

Dataset	WTP Estimate	% Decline
Stratus Original	**$176.78**	**0.0%**
Recode: Alum Without Ban	$149.63	-15.4%
Recode: Certainty	$126.36	-28.5%
Recode: Faster - River	$139.51	-21.1%
Recode: Faster - Lake	$135.95	-23.1%
Recode: Pay Other River/Lake	$98.12	-44.5%
Recode: Income Tax	$116.47	-34.1%
Combined Recodes	**$37.98**	**-78.5%**

[68] The results are fundamentally the same using the ABERS estimator. See Appendix E for results using the ABERS estimator.

Using the recoded data, we estimate WTP for the same subgroups discussed in Section 5.2 of this report. For each of these subgroups, we report both the ABERS and Turnbull estimated results for "all recodes" as well as for each of the six stand alone recodes: "alum without ban", "certainty", "faster – river", "faster – lake", "pay other river/lake", and "income tax." In each instance, WTP declines dramatically in response to the recoding. This demonstrates the extreme fragility of the Stratus results and their instability in light of reclassifications based upon respondent uncertainty or confusion.

Passive Versus Active Users

Dividing each dataset (into those who have previously visited the Illinois River and those who have never visited the Illinois River) confirms that active users have higher WTP than do passive users. The same conclusion holds for Tenkiller Lake. As before, WTP is consistently lower with the recoded dataset than it is with the Stratus-provided dataset, and passive users have lower WTP than active users. These differences are statistically significant for most of the recoded datasets.[69]

Table 5.8 shows the estimated WTP for Active and Passive Users of the Illinois River. The first row presents WTP based on the original Stratus data, calculated with the more appropriate Turnbull estimator. The next rows apply the same procedures to the recoded data to measure how WTP changes. Employing even one of the recodes results in a lower WTP than the estimates using the original Stratus data. Recoding only those respondents who believed that the alum treatments would be done without the ban ("Alum Without Ban" recode) results in the smallest decline in WTP, which nonetheless represents a 14% to 19% decline in WTP for Active and Passive Users, respectively. Each of the other five recodes results in an even greater reduction to the WTP estimates, ranging from 18% to 46% less than the original estimate for both Passive and Active Users. The final row in the table shows the WTP for Passive Users, assuming all six recodes are employed. As this table makes clear, applying all of the recodes results in a WTP estimate between 82% and 92% lower than that estimated by Stratus.

[69] The difference was significant for Alum Without Ban, Certainly, Faster - River, Faster - Lake, and Income Tax.

Table 5.8: Turnbull WTP Estimates for Passive and Active Users of the Illinois River[70]

	Passive IL River		Active IL River	
Dataset	**WTP**	**% Decline**	**WTP**	**% Decline**
Stratus Original	**$154.21**	**0.0%**	**$202.89**	**0.0%**
Recode: Alum Without Ban	$125.27	-18.8%	$174.54	-14.0%
Recode: Certainty	$109.43	-29.0%	$144.16	-28.9%
Recode: Faster - River	$109.17	-29.2%	$109.61	-46.0%
Recode: Faster - Lake	$107.57	-30.2%	$165.71	-18.3%
Recode: Pay Other River/Lake	$89.02	-42.3%	$109.61	-46.0%
Recode: Income Tax	$86.64	-43.8%	$139.52	-31.2%
Combined Recodes	**$26.80**	**-82.6%**	**$17.15**	**-91.5%**

Table 5.9 presents similar results for Passive and Active Users of Lake Tenkiller. Once again, the "Alum Without Ban" recode results in the smallest decline in WTP when compared to the original Stratus estimates; nonetheless, it decreases WTP by between 23% for Active Users and 12% for Passive Users. Indeed, each of the individual recodes reduces the WTP estimate by between 20% and 45%. Accepting all of the recodes results in a WTP that is between 78% and 80% lower than the ABERS estimates using the original Stratus data.

Table 5.9: Turnbull WTP Estimates for Passive and Active Users of Tenkiller Lake[71]

	Passive Lake		Active Lake	
Dataset	**Turnbull WTP**	**% Decline**	**Turnbull WTP**	**% Decline**
Stratus Original	**$142.08**	**0.0%**	**$135.00**	**0.0%**
Recode: Alum Without Ban	$112.81	-20.6%	$138.02	2.2%
Recode: Certainty	$103.03	-27.5%	$148.72	10.2%
Recode: Faster - River	$104.66	-26.3%	$143.68	6.4%
Recode: Faster - Lake	109.81	-22.7%	-	-
Recode: Pay Other River/Lake	$83.55	-41.2%	$112.69	-16.5%
Recode: Income Tax	$86.73	-39.0%	$134.05	-0.7%
Combined Recodes	**$31.92**	**-77.5%**	**$41.04**	**-69.6%**

[70] The results are fundamentally the same using the ABERS estimator. See Appendix D for results using the ABERS estimator.

[71] The results are fundamentally the same using the ABERS estimator. See Appendix D for results using the ABERS estimator.

Effectiveness of Alum Treatment

As with the original data, separating the population by individual perceptions of the effectiveness of alum yields the most dramatic results. The original data reveal differences of over $134 (57%) in the WTP of those who think the treatment will be ineffective or moderately effective and those who think the treatment will be very or extremely effective. While the disparity between the two groups is smaller in terms of a dollar amount ($47), when expressed as a percent the disparity is larger at 84%. Table 5.10 shows the individual effects of each data recode on these two distinct subpopulations, as well as the cumulative change when all six recodes are employed. The individual recodes result in reductions to the WTP of between 15% and 60% when compared with the original Stratus estimates for the combined population. Accepting all of the recodes results in a WTP estimate that is 76% to 91% lower than the ABERS estimate produced with the original Stratus data.

Table 5.10: Turnbull WTP Estimates by Respondents' Belief in Alum Treatment Effectiveness: Not, Slightly, Moderate versus Very or Extremely[72]

Dataset	Not/Slightly/ Moderate	% Decline	Very/ Extremely	% Decline
Stratus Original	**$100.88**	**0.0%**	**$235.29**	**0.0%**
Recode: Alum Without Ban	$82.10	-18.6%	$200.74	-14.7%
Recode: Certainty	$43.48	-56.9%	$181.77	-22.7%
Recode: Faster - River	$62.49	-38.1%	$190.37	-19.1%
Recode: Faster - Lake	$60.34	-40.2%	$186.58	-20.7%
Recode: Pay Other River/Lake	$40.08	-60.3%	$132.72	-43.6%
Recode: Income Tax	$67.42	-33.2%	$147.29	-37.4%
Combined Recodes	**$9.04**	**-91.0%**	**$56.15**	**-76.1%**

State Spending on Pollution

Using the recoded data, we also iterated the distinction between those who want to spend "less" from those who want to spend the "same" on resolving pollution problems. Combining these two subgroups, but using the recoded data, produces the following WTP estimates, which in each case, are lower than Stratus reported for the respondent population at large.

Table 5.11: Turnbull WTP Estimates by Respondents' Belief that State Should Spend Less, or the Same on Pollution (Q7e)[73]

Dataset	WTP Estimate	% Decline
Stratus Original	**$119.96**	**0.0%**
Recode: Alum Without Ban	$106.64	-11.1%
Recode: Certainty	$86.84	-27.6%
Recode: Faster - River	$94.20	-21.5%
Recode: Faster - Lake	$90.79	-24.3%
Recode: Pay Other River/Lake	$65.82	-45.1%
Recode: Income Tax	$62.09	-48.2%
Combined Recodes	**$28.58**	**-76.2%**

5.6 Bid and Income Elasticities Calculated with Recoded Data

As with the original data, the recoded data shows that that the relationship between WTP and respondent income defies economic logic. As income increases, we would expect that WTP would similarly increase, rather than decline. In other words, as in the non-segmented base survey, there is no monotonic relationship between WTP and income. In Table 5.12, we show WTP per income quartile calculated with the recoded data, where "1" represents the highest income quartile and "4" represents the lowest income quartile. Instead of WTP increasing as income increases, WTP increases as income falls with the recoded data. The WTP for the lowest income quartile is $60.56 in the recoded dataset for both the ABERS and Turnbull estimates. In contrast, WTP for the highest income quartile was only $11.46 and $3.05 for the ABERS and Turnbull estimates, respectively. This means that the lowest income group was willing to pay more than five times the amount of the highest income group when the ABERS estimator is employed, or more than 19 times when the Turnbull estimator is used. These results defy economic logic and suggest the infirmity of the Stratus results.

[72] The results are fundamentally the same using the ABERS estimator. See Appendix D for results using the ABERS estimator.

[73] The results are fundamentally the same using the ABERS estimator. See Appendix D for results using the ABERS estimator.

Table 5.12: Willingness-to-Pay by Income Quartile for Recoded Data

Income Quartile	ABERS		Turnbull	
	Original	Recoded	Original	Recoded
4 (lowest)	$187.14	$60.56	$166.10	$60.56
3 (low mid)	$186.94	$55.51	$183.80	$34.27
2 (high mid)	$181.44	$27.21	$175.68	$13.50
1 (highest)	$186.11	$11.46	$173.36	$3.05

In parallel to our elasticity calculations for the base survey, we estimate income elasticities at the mean income for the quartile income groupings. Recall that if the probability of a "yes" vote increases as income rises, then income elasticity will be positive. As found in the base data, although income elasticity is positive for the entire population, when calculated by income groupings, it is not consistently positively related to the probability of voting "Yes."[74] As shown in Table 5.13 for the third quartile of the recoded data, an increase in income is associated with a decrease in the probability of a "yes" vote. Again, these results are inconsistent with economic principles.

Table 5.13: Income Elasticities by Income Quartile

Quartile	Max Income	Original Stratus	Recoded Data
4 (lowest)	$18,000	0.158	0.37
3 (low mid)	$33,000	-0.397	0.07
2 (high mid)	$60,000	0.177	3.576
1 (highest)	$600,001	0.2	5.713

Finally, we reviewed the bid elasticities for the recoded data. Recall that the bid amount was not negatively related to the probability of voting "Yes" in the original data. In other words, for the bid amount of $125, the probability did not decline. For the recoded data, in contrast to the base data, the bid elasticities are now negative at each point. This is more consistent with fundamental economic principles than the results obtained for the base data in the Stratus report. Table 5.14 shows the base bid elasticities for the original Stratus dataset and the recoded datasets.

[74] Elasticities for income quintiles and sextiles are included in Appendix E to demonstrate that this phenomenon is not isolated to income quartiles.

Table 5.14: Bid Elasticities Calculated with Recoded Data

	Bid Elasticity	
Bid	Stratus Data	Recoded Data
$10		
$45	-0.04	0.00
$80	-0.18	-0.51
$125	0.04	-0.14
$205	-0.46	-0.30
$405	-0.22	-0.25

5.7 Results of Scope Test with Recoded Data

Finally, as previously discussed, the effectiveness of the scope test was evaluated using the data as originally coded by Stratus. Using the recoded data, the probability of voting "yes" continues to be much higher for the base survey than for the scope survey. Similarly, estimates of WTP as set forth in Table 5.15 are significantly lower for the recoded data when compared to the original Stratus data.

Table 5.15: Probability of Voting Yes and WTP Estimated With Recoded Data

	Pr(Yes)		ABERS		Turnbull	
	Base	Scope	Base	Scope	Base	Scope
Stratus	0.58	0.42	$184.55	$138.51	$176.78	$138.51
Recoded	0.13	0.07	$37.98	$24.31	$37.98	$6.29

As another comparison between the base and scope surveys, we examined bid elasticities with the results set forth in Table 5.16.

Table 5.16: Bid Elasticities – Stratus vs. Recoded Data

	Main		Scope	
Bid	Stratus Data	Recoded Data	Stratus Data	Recoded Data
$10				
$45	-0.040	-0.001	-0.092	-0.115
$80	-0.183	-0.507	-0.172	-0.924
$125	0.039	-0.143	-0.111	1.891
$205	-0.456	-0.298	-0.304	1.708
$405	-0.221	-0.247	-0.082	-0.478

As with the original Stratus data, the sign for the recoded data changes with the scope survey results. However, in this case, the sign changes and becomes positive for bids of both $125 and $205. This is inconsistent with economic theory, which implies that the all the bid elasticities should be negative. In particular, this result points to large differences in the way people react to bids in the base survey versus the scope test.

To test the validity of our recoded dataset, we administered the same scope analysis employed by Stratus' economic experts on our recoded base and recoded scope data. The econometric tools they used to validate the scope test included an F-test and a logistic regression model. The F-test yields a measure of association and examines the likelihood that voting "yes" is related to being in the base versus the scope study. The logistic regression is used to generate predictions of voting behavior for scope and base participants, conditioning on individual characteristics and a distinguishing variable for whether the respondent was given the scope or base scenario. Our replications of these analyses for recoded datasets corroborate their findings: we find from the F-test that there is a relationship between voting "yes" and being in the base versus the scope study, and we find from the logistic regression that there is a statistically significant, positive effect of the base scenario on the prediction of voting "yes".

The percentage of people who voted "Yes" at each bid are given in Table 5.17 below.

Table 5.17: Recoded Scope Test – Percent of "Yes" Votes

Bid Amount	Recoded base	Recoded scope
$10	19.44%	13.87%
$45	19.40%	8.29%
$80	11.73%	2.30%
$125	10.80%	4.78%
$205	8.73%	10.03%
$405	6.65%	5.37%

5.8 Implications of Pre-Test Survey Data

There is a clear difference in the injury scenario proposed in the pre-tests Stratus conducted and their final base survey. Qualitatively, the injury attributed to phosphorus is far smaller for the base survey than it is in the pre-tests (see Table 5.18). For example, in the first four pre-tests, fish kills were mentioned, the description of algae was extensive, and the extent of fish injury was also extensive. In contrast, the final base survey did not mention fish kills and had a moderate description of algae and moderate fish injury. Table 5.18 shows the bid amounts, including the probability of a "yes" vote for the highest bid, as well as several survey attributes such as the payment vehicle (e.g., tax paid each year for five years versus a onetime tax added to state income tax bill), whether or not the vote was given via a ballot, as well as various measures of injury.

Table 5.18a: Comparison of Pre-Test Surveys and Stratus Main Survey (Bid Amounts)

Component	Jan. 13, 2008	Feb. 4, 2008	Feb. 6, 2008	Pilot 1	Pilot 2	FG 14	Final Base Survey
BidAmount							
10				√	√	√	√
30				√			
45							√
55					√		
60	√	√	√				
65				√		√	
80							√
95					√		
105						√	
115				√			
125							√
150	√	√	√				
155					√		
205							√
245				√			
250	√	√					
375					√		
405							√
500			√				

Table 5.18b: Comparison of Pre-Test Surveys and Stratus Main Survey (Scenario)

Component	Jan. 13, 2008	Feb. 4, 2008	Feb. 6, 2008	Pilot 1	Pilot 2	FG 14	Final Base Survey
Payment Vehicle	Pay tax each year for 5 years	Pay tax each year for 5 years	Pay tax each year for 5 years	Pay tax each year for 5 years	One time tax added to state income tax bill	One time tax added to state income tax bill	One time tax added to state income tax bill
Ballot Envelope	Yes	Yes	Yes	No	No	Yes	No
Fish kills mentioned	Yes	Yes	Yes	Yes	No	No	No
Extent of algae description	Extensive	Extensive	Extensive	Extensive	Moderate	Moderate	Moderate
Types of fish injuries	Extensive	Extensive	Extensive	Extensive	Moderate	Moderate	Moderate
Baseline Years	1960	1960	1960s	1960s	Around 1960	Around 1960	Late 1950s, early 1960s
Restoration Years	River: 50 years	50 years (present with alum 1st at 10 years)	50 years (present with alum 1st at 10 years)	50 years (present with alum 1st at 10 years)	50 years (present with alum 1st at 10 years)	50 years (present with alum 1st at 10 years)	50 years (present with alum 1st at 10 years)
	Lake: 70 years	60 years (present with alum 1st at 20 years)	60 years (present with alum 1st at 20 years)	60 years (present with alum 1st at 20 years)	60 years (present with alum 1st at 20 years)	60 years (present with alum 1st at 20 years)	60 years (present with alum 1st at 20 years)

Despite the smaller injury in the final base survey, the resulting Turnbull WTP estimates show a *higher* WTP, contradicting economic theory. The WTP estimates shown below in Table 5.19 are clearly lower for every pre-test dataset than they are for the final base survey dataset. All but the base surveys conducted on February 6 and for Pilot 2 are lower than the scope version of the final dataset. This suggests that Stratus "improved" its results through ongoing survey design.

Table 5.19: Comparison of WTP and Income Elasticities for Pre-Test Surveys and Stratus Main Survey

Pre-Test Dataset	# of Participants	Turnbull	Difference between base & scope WTP	Income Elasticity (Below median)	Income Elasticity (Above median)
Jan 13, Base	**43**	**$96.93**			
Jan 13, Scope	**37**	**$81.60**	-15.80%	-55.36	0.01
Feb 4, Base	59	$85.36		62.23	22.46
Feb 4, Scope	56	$29.12	-65.90%		
Feb 6, Base	100	$153.77		46.74	-0.61
Feb 6, Scope	**92**	**$118.01**	-23.30%	1.30	
FG14, Base	94	$49.77		0.03	0.27
FG14, Scope	97	$22.75	-54.30%	1.54	
Pilot 1 (Base only)	**152**	**$128.90**	n/a	0.23	0.13
Pilot 2 (Base only)	152	$173.51	n/a	0.12	0.08
Final, Base	1,093	$176.78		0.00	0.00
Final, Scope	544	$138.51	-21.60%	-0.001	-0.006

We also find no consistent relationship between income and WTP using the logit specification provided in the Stratus report. We divide household incomes into two groups, below and above the median household income of each set of respondents, and report the income elasticity calculated for each group. None of the elasticity estimates is statistically significant, indicating that there may be no relationship between household income and response. This runs counter to simple economic logic and prior studies regarding WTP for preserving or improving environmental resources.

The comparison of pre-test and final survey results also highlights the serious problem inherent in contingent valuation based on bid structure. Since nonparametric WTP estimators use weighted averages of the bids, any bid above $405 added to the bid structure would have increased the mean WTP in all cases, unless the bid was so high that no respondents would accept it. However, we are aware of no CV studies with a significant nonuse component where the authors have offered a bid high enough that the number of "yes" responses approaches zero. With the evidence that some respondents will say "yes" to almost any bid,[75] it is possible to predetermine the mean

[75] Carson, et al. (1992) study the Exxon Valdez oil spill and find 34% of respondents say "yes" at $120. Carson, et al. (1996) study the Southern California Bight and find that 25% say "yes" at $215. Brown, et al. (1996) study the Grand Canyon and find 33% say "yes" at $50. Desvousges, et al. (1993) study of oil-spill response finds 30%-38% say "yes" at $1,000. McFadden and Leonard (1993) study the Selway wilderness area and find 26 % say "yes" at $2,000.

WTP just by the selection of the highest bid. This frequent phenomenon in contingent valuation has been interpreted as indicating "yea-saying" responses. Yea-saying artificially increases the proportion of people who respond "yes" at any given bid, magnifying the resulting hypothetical bias.

Focusing our attention on the base survey, we see that 34.17% of respondents offered the highest bid of $405 responded "yes." Because the CV survey only allows "yes" or "no" answers, there is very little information about the nature of the true WTP for these respondents. Testing for the sensitivity of estimated WTP to yea-saying, we assume that everyone who said "yes" to this bid either has an outlier WTP or would have responded "no" given more time to dwell on their preferences. With this assumption, we estimate a Turnbull WTP of $108.42, with a 95% confidence interval of $93.76 to $123.08. This represents a decrease of 38.67% from the Turnbull WTP of $176.78, estimated from the base survey and assuming no degree of nay-saying. With this extreme sensitivity, the authors must provide some serious justification for their implicit assumption of yea-saying not causing substantive hypothetical biases.

5.9 References

Alberini, A., Rosato, P., Longo, A., Zanatta, V. 2004. "Information and Willingness to Pay in a Contingent Valuation Study: The Value of S. Erasmo in the Lagoon of Venice." Nota Di Lavoro 19.2004.

Bateman, I.J., and Langford, I.H. 1997. "Non-users' Willingness to Pay for a National Park: An Application and Critique of the Contingent Valuation Method." *Regional Studies* 31.6: 571-582.

Brown, Thomas C., Patricia A. Champ, Richard C. Bishop and Daniel W. McCollum. 1996. "Which Response Format Reveals the Truth about Donations to a Public Good?" *Land Economics* 72(2): pp.152-166.

Carson, Richard T, Robert C. Mitchell, Michael Hanemann, Raymond J. Kopp, Stanley Presser, and Paul A. Ruud. 2003. "Contingent Valuation and Lost Passive Use: Damages from the Exxon Valdez Oil Spill." *Environmental and Resource Economics* 25(3): 157-286.

Carson, Richard T., Michael B. Conaway, W. Michael Hanemann, Jon A. Krosnick, Kerry M. Martin, Donald R. McCubbin, Robert Cameron Mitchell, and Stanley Presser. 1996. *The Value of Preventing Oil Spills to Natural Resources Along California's Central Coast.* Report prepared for the California Department of Justice, the Office of Oil Spill Prevention and Response in the California

Department of Fish and Game, and the Damage Assessment Center of the National Oceanic and Atmospheric Administration. San Diego, CA: Natural Resource Damage Assessment, Inc.

Desvousges, William H., F. Reed Johnson, Richard W. Dunford, Kevin J. Boyle, Sara P. Hudson, and K. Nicole Wilson. 1993. "Measuring Natural Resource Damages With Contingent Valuation: Tests of Validity and Reliability." In *Contingent Valuation: A Critical Assessment*, Jerry A. Hausman, ed. Amsterdam: Elsevier Science Publishers B.V.

Fisher, T.C.G., and Waschik, R.G. 2002. "Chapter 11: Government regulation of industries." In *Managerial Economics: A Game Theoretic Approach.* Lodong, New York: Routledge.

Haab, Timothy C. and Kenneth E. McConnell. 2003. *Valuing Environmental and Natural Resources: The Econometrics of Non-market Valuation.* Edward Elgar Publishing.

Jacobsen, B.R., and Hanley, N. 2008. "Are there income effects on global willingness to pay for biodiversity conservation?" Submitted to *Environmental and Resource Economics.*

Kriström, Bengt and Pere Riera. 1996. "Is the Income Elasticity of Environmental Improvements Less Than One?" *Environmental and Resource Economics* 7(1) 45-55.

Le Goffe, Ph. 1995. "The Benefits of Improvements in Coastal Water Quality: A Contingent Approach," *Journal of Environmental Management* 45: 305-317

McFadden, Daniel, and Gregory K. Leonard. 1993. "Issues in the Contingent Valuation of Environmental Goods: Methodologies for Data Collection and Analysis." In *Contingent Valuation: A Critical Assessment*, J. A. Hausman, ed. Amsterdam: Elsevier Science Publishers B.V.

Organisation for Economic Co-operation and Development. 2001. *OECD Environmental Outlook*, OECD Publishing.

Scheierling, S. M., J. B. Loomis, and R. A. Young. 2006. "Irrigation water demand: A meta-analysis of price elasticities." *Water Resour. Res.* 42.

Schläpfer, F., "Survey protocol and income effects in the contingent valuation of public goods: A meta-analysis," *Ecological Economics* 57 (2006): 415-429.

Soderqvist, Tore, and Therese Lindahl. 2003. "Wetland Creation: Socio-Economic and Institutional Conditions for Collective Action." In *Managing wetlands: An ecological economics approach,* Turner, R. Kelly, van-den-Bergh, Jeroen C. J.M., Brouwer, Roy, eds. Cheltenham, UK and Northampton, Mass, pp: 223-49.

6. AGGREGATION OF DAMAGES

To aggregate future damages, the plaintiffs' experts multiply the resulting WTP from the CV survey by the number of households in the counties sampled during the survey process. This aggregation essentially assumes that all households in the 63 counties hold either (or both) use values or nonuse values for the aesthetics of the Illinois River System and Tenkiller Lake. However, from an economic perspective, this number of affected households is an unproven assertion advanced by the plaintiffs' experts.

The economics literature has long recognized that only some individuals or households have economic standing with respect to the quantification of societal benefits (Whittington and MacRae 1986). However, at the time of the NOAA Panel report, the issue of "who counts" had not yet been raised by economists. Subsequent to the NOAA Panel, Smith (1993) and Dunford, Johnson, and West (1997) extend the logic of economic standing to recreation services provided by natural resource services, revealing that who counts is not a foregone conclusion from an economic perspective. For use values, observed behaviors guide the determination of the extent of the market.

With respect to nonuse values (or total values because they include nonuse values), reliance on geopolitical boundaries results in an arbitrary and unsupported determination of who counts. This arises because "no simple rules define who holds these values" (Bateman 2000). Thus, the convenience of using geopolitical boundaries results in inaccurate damage estimates. Hanley, Schläpfer, and Spurgeon (2003) conclude that:

> "[e]rrors made in estimating the number of users and non-users effected [sic] by an environmental change can easily swamp errors in estimates of per-person Willingness to Pay (WTP) when aggregate values are calculated" (p. 297).

Bateman, et al. (2006) provide other empirical examples of the overestimates produced by reliance on the geopolitical boundaries.

From an economic perspective, having standing for nonuse damages requires that an individual be in a position to experience an economic welfare loss from a specific natural resource injury. Thus, logically, the natural resource service must be a component of the individual's utility or well-being. If that individual has no knowledge of a natural resource, then it cannot affect his or her well-being. Knowledge of a natural resource is thus a necessary, but not sufficient, condition for a welfare loss. As noted earlier, the CV survey respondents' awareness of water quality was created within the Stratus survey and was based on biased and misleading information. This further limits the ability to draw any kind of scientific conclusion about what other households in Oklahoma would have thought about water quality in the Illinois River System and Tenkiller Lake.

In addition, the individual must perceive a difference in the quality (or quantity) of a natural resource service in order to experience an economic welfare loss. Bockstael, et al. (2000) emphasize that measuring the value of a natural resources depends is relative to current conditions. Thus, noticing a change in the resource satisfies both the necessary and sufficient conditions.

Johnson, et al. (2001) develop a conceptual model for nonuse values that establishes the roles of knowledge and awareness. Their theoretical model follows the work of Kaldor and Hicks and relies upon the well-established economic principle of Pareto improvement for welfare measurements. They measure knowledge of a distressed river system in the Northeastern United States by conducting a knowledge survey of households within 400 miles of the river. Like the 2006 survey work by Stratus, they ask about respondents' knowledge without prompting them or informing them. Their case study demonstrates that within the state boundaries, less than 50 percent of the households within the state had knowledge of the river. When evaluating the awareness criterion of households surveyed, even fewer households (about 15 percent) were aware of the environmental changes. Thus, state residency is not a reliable proxy for knowledge or awareness of the resource, further reducing any rationale for the Stratus decision to multiply their survey results by the number of households in most of Oklahoma.

The Stratus CV survey does not provide the opportunity to assess knowledge or awareness separate from the information provided by the interviewer. Respondents are first told that the Illinois River is a scenic river. After being told that it is a scenic river, about one-third of the respondents claim that they knew about its scenic status prior to the interview. Similarly, respondents are also told about the alleged algae conditions. After being told by the interviewers that water clarity is worse now than it was in 1960, about one-third of the respondents claim that they had prior knowledge of the change. Following the lead of Carson, et al. (1994), the Stratus CV survey "constructs" nonuse values. Kontoleon, Macrory, and Swanson (2002) state in regard to the construction of nonuse values:

> "Respondents in CV studies that have not (endogenously) acquired such information nevertheless receive (exogenous) information from the study itself…The usefulness of the estimated values from such individuals for damage assessment is questionable. [Nonuse values] do not exist independent of individual perception. Hence, losses in nonuse values require some prior knowledge" (pp. 197-198).

Thus, the combination of nonresponse bias in the CV survey results, the inability to of the CV survey to demonstrate prior knowledge of the affected resources, and the lack of any empirical rationale for the number of affected households renders the Stratus estimate of the number of households invalid. Given that this number is a large part of the total damage estimate generated by Stratus, the total estimate itself is invalid.

6.1 References

Bateman, I.J., B.H. Daya, S. Georgioua, and I. Lake. 2006. "The Aggregation of Environmental Benefit Values: Welfare Measures, Distance Decay and Total WTP." *Ecological Economics* 60: 450–60.

Bateman, Ian J., Andrew P. Jones, Naohito Nishikawa, and Roy Brouwer. 2000. "Benefits Transfer in Theory and Practice: A Review." CSERGE Working Paper GEC 2000-25. Bishop0000339_Bateman.et.al.2000pdf.pdf.

Bockstael, Nancy E., A. Myrick Freeman, III, Raymond J. Kopp, Paul R. Portney, and V. Kerry Smith. 2000. "On Measuring Economic Values for Nature." *Environmental Science and Technology*. 34(8): 1384-1389.

Carson, Richard T., W. Michael Hanemann, Raymond J. Kopp, Jon A. Krosnick, Robert C. Mitchell, Stanley Presser, Paul A. Ruud, and V. Kerry Smith. 1994. *Prospective Interim Lost Use Value Due to DDT and PCB Contamination in the Southern California Bight.* Report prepared for the National Oceanic and Atmospheric Administration. La Jolla, CA: Natural Resource Damage Assessment, Inc. September 30.

Dunford, R. W., Johnson, F. R., and West, E. S. 1997. "Whose Losses Count In Natural Resources Damages?" *Contemporary Economic Policy* 15(4): 77(11).

Hanley, N., F. Schläpfer, and J. Spurgeon. 2003. Aggregating the Benefits of Environmental Improvements: Distance-Decay Functions for Use and Non-use Values. *Journal of Environmental Management* 68: 297–304.

Johnson, R.F., R.W. Dunford, W.H. Desvousges, and M.R. Banzhaf. 2001. The Role of Knowledge in Assessing Nonuse Damages: a Case Study of the Lower Passaic River. *Growth and Change* 32: 43–68.

Kontoleon, Andreas, Richard Macrory, and Timothy Swanson. "Individual preference-based values and environmental decision making: should valuation have its day in court?" *Research in Law and Economics* 20:179–216.

Smith, V.K. 1993. "Welfare Effects, Omitted Variables, and the Extent of the Market." *Land Economics* 69(2): 121–31.

Whittington, D. and D. MacRae, Jr. 1986. "The issue of standing in cost-benefit analysis." *Journal of Policy Analysis and Management* 5:665.

7. CRITIQUE OF PAST DAMAGES STUDY

The damage estimate from the Stratus CV study addresses only the alleged future losses in services provided by the Illinois River System and Tenkiller Lake. The damage estimate from the Stratus CV survey does not represent alleged past losses in services provided by the lake and river system. In order to generate additional damages that reflect the alleged past losses in services, the Stratus team pro-rates the WTP results from the CV survey and applies that pro-rated amount to past years (1981 to 2008). This section discusses the serious flaws with that approach, which render the estimate of past damages unreliable.

7.1 The methodology is not consistent with a benefits-transfer approach.

Stratus presents its pro-rating process as a benefits-transfer. A benefits-transfer, as customarily discussed in the literature, uses existing WTP results, based on an original data from one geographic area, and applies them to another geographic area. These transfers are most often used in policy applications, where collecting original data is not financially feasible. The literature on benefits transfer also clearly notes the limitations of the method. For example, for valuing ecosystem services, such as those purportedly measured by the Stratus CV survey, the National Research Council (NRC) (2005) indicated that "benefits transfer is generally considered a "second best" valuation approach by economists" (p.124). The NRC adds that such transfer should be viewed with caution and done according to strict guidelines. Smith (1996) indicates that we have not done much research on benefits transfer, and instead have merely performed such transfers.

The existing literature demonstrates the notion that a benefits-transfer is a spatial concept. For example, the seminal studies on benefits-transfer published in the early 1990s reveal an explicit focus on transferring WTP from a "study site" to a "policy site" (Brookshire and Neill 1992; Smith 1992; Desvousges, Naughton, and Parsons 1992; McConnell 1992; Boyle and Bergstrom 1992). More recent literature confirms the defining spatial feature (Desvousges, Johnson, and Banzhaf 1998; EVRI). For example, Zandersen, Termansen, and Jensen (2007) offer the following definition:

> "Benefits transfers are based on sites where monetary valuation has already been carried out (policy sites) and transferred to new, unstudied sites (study sites)…Benefit transfers have traditionally been carried out over space from one geographical location to another" (p. 412).

Because the Stratus approach is not consistent with the established literature on benefits-transfer, the methodology used by Stratus is neither well-established nor generally accepted by the economics profession.

7.2 Applying values backwards in time is not reliable.

To our knowledge, the literature on benefits transfer contains no references to studies that extrapolate damages backward in time. In fact, only a handful of studies have evaluated the temporal aspects of applying WTP forward in time, and none of these studies concludes that doing so provides reliable estimates of WTP. Loomis (1989) finds evidence that WTP values may be relatively stable over short periods of time (nine months) when the determinants of WTP stay constant. Downing and Ozuna (1996) investigate the reliability of applying WTP values three years in the future. They conclude that applying values over time is not reliable. Zandersen, Termansen, and Jensen (2007) test the accuracy of a forward application of recreational values over a period of 20 years for 52 forests in Denmark. They find error rates ranging from 25 to nearly 300 percent over the 20-year span.

In contrast to this handful of studies, none of which concludes that a forward application is reliable, the Stratus methodology take a current WTP estimates and applies it ***backwards*** for 28 years. There is no literature to support the reliability of either the backwards application or the length of time. As the above literature shows, even when forecast for periods as short as a few years, the results have not been reliable. Accordingly, the Stratus methodology, which is forecast backwards for more than twenty five years, is not reliable.

7.3 Stratus fails to demonstrate that preferences for improved water quality are constant.

One reason that the temporal application of WTP estimates is not reliable is because the preferences that govern true WTP values are not constant. Preferences for natural resource services conform to economic principles. (See Smith, Van Houtven, and Pattanayak 1999.) They will reflect the dynamic nature of the quantity and quality of substitute services, as well as budget constraints. Because these features change over time, preferences change over time. Thus, WTP values should not be expected to be constant over time.

In an attempt to address this point, Stratus relies on results from the General Social Survey (GSS). This survey has been conducted annually for decades and evaluates social trends. Since the 1970s, the survey has asked two questions about environmental spending. Stratus reviews responses to these questions over time and concludes that there has been "no material change in attitudes towards spending on the environment" between 1980 and the present. Based on this survey, Stratus proceeds with its backwards application of WTP values.

This logic is flawed for at least two reasons. First, the population surveyed in the GSS study encompasses three states beyond Oklahoma: Arkansas, Louisiana, and Texas. The sheer size of Texas will dominate the GSS results. Specifically, Texas has at least four times the number of households that Oklahoma has (US Census 2009). Although the GSS survey may accurately reflect the environmental spending preferences of the four-state region, Stratus cannot demonstrate that it reflects the preferences for 63 Oklahoma counties.

In addition, according to the construct validity model developed by Stratus in its CV report, attitudes on environmental spending is only one of a myriad of beliefs and opinions that may have influenced how the CV respondents voted. Primary influences included a number of study-specific opinions and beliefs, such as the speed of natural recovery, the seriousness of the algae issue,[76] the effectiveness of the alum program,

[76] Specifically, Stratus' own 2006 telephone survey (see Section 2 above) demonstrates that in 2006 only a small percentage of Oklahoma residents would agree that the algae was a serious issue.

the expectations about the cost of the alum program, and many others. Without demonstrating that all of these opinions and beliefs have been constant from 1980 to 2008, Stratus cannot reliably apply the WTP estimates from the CV survey backwards in time. They offer little or evidence to support either of these critical assumptions.

Finally, it is important to consider that the GSS survey asks people only about a general attitude toward the environment. Such a general attitude is likely to have little predictive ability in explaining people's actual trade-offs that would have made over the twenty year period. It does not present provide any indication that people in the 1980's would actually have spent the same on improving the environment as people today. Nor does it provide any specific support to the potential reliability of the estimates in the Stratus CV survey.

7.4 The approach is not sufficiently reliable for litigation.

Even if the Stratus methodology were a benefits-transfer, which it is not, the literature reveals that benefits-transfer is not sufficiently reliable for litigation purposes. Brouwer and Spaninks (1999) provide "a disappointing result of how [benefit-transfer] fails even when study sites are close and the environmental good is identical" (Bishop undated). The authors compare the findings of two CV studies carried out in the Netherlands shortly after each other with regard to agricultural wildlife management on Dutch peat meadow land. Both studies concentrated on the same type of environmental good in similar areas. However, when transferring WTP from one site to another, the authors rejected the validity of transferring benefit functions.

Finally, the literature reveals the professional judgment of natural resource economists that it is not sufficiently reliable for litigation. For example, Bergstrom and Taylor (2006) state that benefits-transfer may be sufficiently accurate "for applications requiring low to moderate accuracy (e.g., screening, minor policy decisions). For applications requiring moderate to high accuracy (e.g., litigation, major policy decisions), primary data studies will probably still be preferred" (p. 359). Bergstrom and De Civita (1999) note several errors that arise with benefits transfer and reveal that "if benefits transfer is used as a basis for determining just compensation in the context of

natural resource damage litigation, the costs of a wrong decision to individuals and society could be quite high" (p. 83). Navrud (2001) notes that errors associated with uncertainty in benefits transfer can be quite large. He contends that benefits transfer should be applied to uses of environmental valuation where the demand for accuracy is not too high. "More caution should be exercised in using transferred values… in natural resource damage assessments" (p. 72).

7.5 The Past Damages report is not based on a valid study.

Even if the Stratus methodology were a benefits-transfer, which it is not, it fails to meet on the long-established criteria for a valid transfer. Scientific soundness refers to the overall quality of a study and is widely recognized as a primary criterion for applying the results from one study to another situation (Brookshire and Neill 1992; Smith 1992; Desvousges, Naughton, and Parsons 1992; McConnell 1992; Boyle and Bergstrom 1992; Desvousges, Johnson, and Banzhaf 1998; EPA 2000). The quality encompasses all aspects of a study, such as the data, the methodology, the survey protocols, and the analysis technique. This criterion effectively asks whether the original study is sufficiently sound science. If the results were not based on reliable data, rigorous protocols, and valid analyses, then the results are not reliable and should not be used in a benefits transfer.

The past damages monetary claim made by the plaintiffs depends critically on the Stratus CV Survey. Sections 4 and 5 of this report documents the extent of hypothetical bias, nonresponse bias, the lack of balance in the survey questionnaire, the absence of validity in the CV results, and the consistent upward bias in the estimation protocols. For all of the reasons documented above, the Stratus CV results are not scientifically valid. Therefore, the benefits transfer of the CV results to past damages renders the past damages estimate invalid and unreliable.

7.6 The Past Damages report relies on faulty scientific assumptions.

The validity of certain scientific opinions enters into the calculation of the past damages. Because the scientists working for the plaintiffs believe that the average annual injury is comparable between 1981-2008 and 2009-2063, the Stratus team pro-

rates the CV results to the number of past years. However, an important implicit assumption in this pro-rating scheme is that the presence of the poultry industry has been constant since 1981. On the contrary, the inventory of meat-type chickens in the relevant Oklahoma counties in 2007 was twice the inventory in 1987. (See Census of Agriculture 1987,1992,1997,2002, and 2007.) Moreover, the Stratus analysis fails to account for changes in water quality conditions in watershed since 1981 that are influenced by population growth and the associated impacts on water quality through increased numbers of septic systems and more waste water treatment plants, among other factors. The failure to account for other factors means that the Stratus approach to estimating past damages results in an overstatement of past damages.

7.7 The Past Damages assumptions about compound interest are flawed.

Compound interest plays a crucial role in the Stratus past damages calculations. Specifically, of the total past damages demand, nearly two-thirds of it is attributable to compound interest. As economists and not lawyers, we do comment on whether it is within the court's discretion to award compound interest in legal matters. However, we note that awarding compound interest in this case does not reflect actual funds that were lost from the State's coffers. The damages claimed by plaintiffs' consist largely of respondents' nonuse, or passive use values for a hypothetical restoration program that is neither safe nor effective and to prevent a highly biased set of injuries. Moreover, the earlier Stratus study results from actual users depicted a very different picture of water quality in the Illinois River and Tenkiller Lake. Thus, these purported losses were not experienced by people who necessarily have visited the area—in fact, half have not. These purported losses are not based on the loss of some type of productive asset that the citizens could have invested to earn interest. Thus, there is no economic basis to award compound interest for these hypothetical losses.[77]

7.8 References

Bergstrom, J. C. and L. O. Taylor. 2006. "Using meta-analysis for benefits transfer: Theory and practice." *Ecological Economics* 60(2): 351-360.

[77] CVs for William H. Desvousges, Ph.D. and Gordon C. Rausser, Ph.D. are in Appendix F.

Bergstrom, J. and P. De Civita. 1999. "Status of benefits transfer in the United States and Canada: A review." *Canadian Journal of Agricultural Economics* 47(1): 79-87.

Bishop undated. Bishop0000312_Benefit.Transfer.review_04-21_eh.doc

Boyle, K.J., and J.C. Bergstrom. 1992. "Benefit Transfer Studies: Myths, Pragmatism, and Idealism." *Water Resources Research* 28(3):657–663.

Brookshire, D.S., and H.R. Neill. 1992. "Benefit Transfers: Conceptual and Empirical Issues." *Water Resources Research* 28(3):651–655.

Brouwer, R. and F. Spaninks. 1999. "The validity of transferring environmental benefits: Further empirical testing." *Environmental and Resource Economics* 14(1): 95-117.

Desvousges, W.H., F.R. Johnson, and H.S. Banzhaf. 1998. Environmental Policy Analysis With Limited Information: Principles and Applications to the Transfer Method. Cheltenham, UK: Edward Elgar.

Desvousges, W.H., M.C. Naughton, and G.R. Parsons. 1992. "Benefit Transfer: Conceptual Problems in Estimating Water Quality Benefits Using Existing Studies." *Water Resources Research* 28(3):675–683.

Downing, M. and T. Ozuna. 1996. "Testing the Reliability of the Benefit Function Transfer Approach." *Journal of Environmental Economics and Management* 30(3): 316-22.

Environmental Valuation Reference Inventory. http://www.evri.ca/

Loomis, John B. 1989. "Test-retest Reliability of the Contingent Valuation Method: A Comparison of General Population and Visitor Responses." *American Journal of Agricultural Economics* 1: 76-84.

McConnell, K.E. 1992. "Model Building and Judgment: Implications for Benefit Transfers with Travel Cost Models." *Water Resources Research* 28(3):695–700.

National Research Council. 2005. *Valuing Ecosystem Services Toward Better Environmental Decision–Making.* The National Academies Press. Washington, D.C.

Navrud, S. 2001. "Comparing valuation exercises in Europe and the United States – challenges for benefit transfer and some policy implications." In: *Valuation of Biodiversity Benefits: Selected Studies*. Biller, D. and Bark, R. (eds). OECD, Paris, pp. 63-77.

Smith, V.K., Van Houtven, G., Pattanatak, S. 1999. "Benefit Transfer as Preference Calibration." Resources for the Future, Discussion Paper 99-36.

Smith, V.K. 1992. "On Separating Defensible Benefit Transfers from Smoke and Mirrors." *Water Resources Research* 28(3):685–694.

U.S. Census Bureau. 2009. "QuickFacts." Accessed March 11, 2009.

U.S. Census of Agriculture. 1987,1992,1997,2002 and 2007

U.S. Environmental Protection Agency. 2000. *Guidelines for Preparing Economic Analyses.* Report Number 240-R-00-003. Washington, DC: USEPA. September.

Zandersen. M., Termansen, M. and F. S. Jensen. 2007. "Testing benefits transfer of forest recreation values over a twenty-year time horizon." *Land Economics* 83(3): 412-440.

Appendix A
Literature Review On Hypothetical Bias

The Empirical Literature Confirms Hypothetical Bias in CV Results

To evaluate how the hypothetical nature of the questions affects CV results, researchers have performed several experiments to test for hypothetical bias. Generally, these studies find hypothetical bias. Vossler, et al. (2003) reveal that a majority of the studies find significant response differences in hypothetical and real situations. Murphy and Stevens (2004) note that the literature shows hypothetical bias across a wide variety of CV approaches. Johnston (2006) concurs: "Most research finds significant divergence between stated and actual behaviors" (p. 469).

Following Vossler and Kerkvliet (2003), these studies fall into four groups. The first group of studies tested the difference between actual payments for private goods and stated CV payments for the same private goods (Bishop and Heberlein 1979; Dickie, Fisher, and Gerking 1987; List and Shogren 1998; Cummings, Harrison, and Rutström 1995; Berrens and Adams 1998). The second group of studies evaluated the difference in stated CV payment and revealed actual WTP for public goods for which observed behaviors are available (Knetsch and Davis 1966; Brookshire, et al. 1982; Loomis, Creel and Park 1991; Shabman and Stephenson 1996). These two groups of studies reflect use values for natural resource services and generally demonstrate hypothetical bias.

The third group of studies developed simulated market experiments to test whether CV values are comparable to the amount respondents would really pay if an actual market existed. This synthetic form of external validation involves comparing CV values to actual cash payments from a simulated market for the same commodity (Kealy, Montgomery, Dovidio 1990; Seip and Strand 1992; Bohm 1992; Duffield and Patterson 1992; Brown et al. 1996; Champ et al. 1997). For example, Duffield and Patterson (1992) compare stated and actual WTP for maintenance of instream water flows in Montana and find that CV values exceed actual payments by a factor of 4 for residents and a factor of 3 for nonresidents. Brown et al. (1996) elicited WTP for a road-removal program on the North Rim of the Grand Canyon. The results show that mean stated WTP was four to six times the mean actual WTP, with the means being statistically different.

The last group of studies contains studies that use the referendum format to elicit WTP values for various types of goods, both private and public. In a number of these studies, the CV referenda did not reflect an actual referenda, much like the CV study that Stratus has conducted for this litigation (Cummings et al 1997; Bjornstad et al. 1997; Taylor 1998; Cummings and Taylor 1999; Taylor et al. 2001; Brown et al. 2003; Landry and List 2007; Burton et al. 2007; Carson, Groves, and List 2008). However, an actual group payment was part of the study design. Overall, these studies reveal that hypothetical bias persists, even when the referendum format is used.

A subset of the referenda studies has compared CV results from a simulated referendum for a public good to voting results from an actual referendum for the identical public good (Carson, Hannemann, and Mitchell 1986; Shabman and Stephenson 1996; Champ and Brown 1997; Vossler and Kerkvliet 2003; Vossler et al. 2003; Schläpfer Roschwitz, and Hanley 2004; Johnston 2006). Bishop (undated) contends that these studies represent the best test of hypothetical bias for the Stratus CV study and that such studies do not, generally, reflect hypothetical bias. However, a closer examination of these studies contradicts those conclusions.

Table 3.5 summarizes this set of studies. In addition to the fact that the Stratus CV survey does not reflect an actual referendum, there are three additional features of these studies that are relevant to a discussion of hypothetical bias in the Stratus CV survey. The first feature is the nature of the public good, shown in the second column of Table 3.5. The studies that do not exhibit hypothetical bias involve public services actually used by the voters. These referenda asked voters to approve bonds or other funding for the construction of sewage treatment plants, public road maintenance and improvements, public water supply provision, and river front park improvements. The one exception to this conclusion is the Shabman and Stephenson (1996) study of flood protection projects.

The commodities depicted in the two studies that clearly demonstrate hypothetical bias are open space preservation and rural landscape protection. While some voters may use open spaces and directly benefit from some rural landscape protection, other voters will not use these types of natural resource services. When the

commodities of the hypothetical referenda studies are examined, the majority (but admittedly not all) of them are also largely nonuse commodities. Thus, the use/nonuse distinction likely explains at least part of the findings on hypothetical bias in referenda studies. Cameron and Englin (1997), Blamey, et al. (2001), Johnston, et al. (1995), and Johnston (2006) all demonstrate that first-hand experience or familiarity with the good leads to a closer correspondence between stated intentions and actual behaviors.

Table A.1: Empirical Studies on Actual Referenda

Study	Referendum	Familiarity and Salience of Proposed Project	Evidence of Hypothetical Bias?
Carson, Hanneman and Mitchell (1987)	Construction of sewage treatment plants in California in 1984	High familiarity and salience. No additional information provided in the survey.	No, only if undecided responses are recoded as no
Johnston 2006	Provision of public water supply to Village of North Scituate, RI in 2001	High familiarity and salience. No additional information provided in the survey.	No
Vossler and Kervliet 2003	Riverfront park improvements in downtown Corvallis, OR in 1998	High familiarity and salience. Community had studied the issue for 6 years. No additional information provided in the survey.	No*
Vossler, et al. (2003)	Purchase of open space near Corvallis, OR in 1995	High familiarity and salience. Most discussed item on the ballot that year. No additional information provided in the survey.	No, only if undecided responses are recoded as no
Schläpfer, Roschwitz, and Hanley (2004)	Improved protection of rural landscape near Zurich, Switzerland in 1996	Some familiarity and salience. The CV survey provided substantial information.	Yes
Shabman and Stephenson (1996)	Flood protection project in Roanoke, VA in 1989	High familiarity and salience. Flooding two years earlier caused $200 million in property damage. The CV survey provided substantial information.	Yes

Study	Referendum	Familiarity and Salience of Proposed Project	Evidence of Hypothetical Bias?
Champ and Brown (1997)	Use of budget surplus for road maintenance in Fort Collins, CO in 1996	High familiarity and salience. No additional information provided in the survey.	No
* However, the study was also designed to test the treatment of undecided votes. If undecided votes are re-coded as votes against, then there is a statistical difference between the actual vote and the survey results.			

The Stratus CV study has elements of both use and nonuse. Approximately 50 percent of the respondents answered "yes" when asked if they had ever visited the Illinois River or Tenkiller Lake (Tables D.14 and D.15). But less than 20 percent of the base version respondents indicate that they have visited in the last three years. Moreover, the open-ended responses from respondents who voted for the program indicate that respondents were thinking of their children, grandchildren, or others when they voted for the program (Table D.89). In light of the hypothetical bias results in referendum studies for commodities that have a nonuse component, hypothetical bias remains a fatal flaw in the Stratus CV study.

The second feature of the actual referenda studies that merits discussion is the salience and familiarity of the good to survey respondents. Certainly, this feature is related to the use values aspects identified above. However, what is an important extension of that concept is the amount of information provided to the CV survey respondents in advance of their votes in the survey. For the majority of these studies, the survey designers did not have to provide information about the issues to the respondents. In fact, Johnston (2006) believes that this lack of additional information is one of the reasons that his study does not exhibit hypothetical bias. For most of these studies, the survey respondents had access to information about the ballot issue from a variety of sources and viewpoints. The two studies that did provide substantial information to the respondents exhibit hypothetical bias. This feature is relevant to the evaluation of hypothetical bias for the Stratus CV survey. Recall that the earlier Stratus surveys from 2006 revealed minimal awareness of the algae conditions. Thus, the Stratus CV survey included a substantial amount of information in order to "educate" the respondents prior to their hypothetical votes. Providing so much information to

respondents is a symptom of the lack of salience and a corresponding likelihood of hypothetical bias in the Stratus CV study.

The last feature of the actual referendum studies that is relevant to a discussion of hypothetical bias is the treatment of the undecided voters in the CV survey. Two of the studies that do not find hypothetical bias do so only because they treat the undecided votes as votes against. In addition, Vossler and Kerkvliet's (2003) study design includes a separate element to test for the treatment of undecided votes. They find that there is no clear evidence that undecided votes should be treated as votes against and that doing so results in statistical differences. Vossler et al. (2003) conclude that it is an open question whether undecided votes should be recoded as votes against. Wang (1997) reasons that "common sense suggests that if a respondents is answering truthfully, a DK [don't know/not sure/would not vote] response is not the same as no" (p. 220).

In an actual referendum, the undecided votes would not be counted, either because the voters did not go to the polls or because they did not make an explicit choice on their ballot. Treating the undecided votes as votes against is particularly important when predicting the WTP for the commodity. Specifically, without such an adjustment of the undecided votes, the survey results over-predict both the percentage of votes for the proposition and the WTP for the commodity at issue. This finding is pertinent to the Stratus CV survey because the Stratus study did not allow the no-vote option. Perhaps this tendency to overestimate WTP was the motivation behind the NOAA Panel's recommendation for a no-vote option. Had the Stratus study included a no-vote option, it likely would have found similar patterns.

Thus, hypothetical bias is prevalent in empirical studies. Most studies that claim to find no hypothetical bias depend on a manipulation of the undecided responses. The three studies that do not find hypothetical bias, without manipulation of the undecided responses, reflect use value goods without significant information dosing in the survey questionnaire. The weight of the evidence suggests that hypothetical bias is likely present in the Stratus CV study.

A.1 References

Bishop, Richard C. and Thomas A. Heberlein. "Measuring Values of Extramarket Goods: Are Indirect Measures Biased?" *American Journal of Agricultural Economics* 61(5): 926-930.

Bishop undated. Notes on Hypothetical Bias. Bishop0001271_Notes on Hypothetical Bias.doc

Blamey, R., J. Bennett, J.J. Louviere, and M. Morrison. 2001. "Green product choice." In: J. Bennett, R. Blamey (Eds.), *The Choice Modelling Approach to Environmental Valuation*, Edward Elgar, Northampton, MA.

Cameron, T.A. and J. Englin. 1997. "Respondent experience and contingent valuation of environmental goods." *J. Environ. Econom. Manage.* 33(3): 296–313.

Bjornstad, David, Ronald Cummings, and Laura Osborne. 1997. "A Learning Design for Reducing Hypothetical Bias in the Contingent Valuation Method." *Environmental and Resource Economics* 10: 207–221.

Bohm, P. 1992. "CVM Spells Responses to Hypothetical Questions." Working paper. Stockholm: University of Stockholm, Department of Economics.

Brookshire, D.S., M. Thayer, W.D. Schulze, and R. d'Arge. 1982. "Valuing public goods: a comparison of the survey and hedonic approaches." *Amer. Econom. Rev.* 72(1):165–177.

Brown, Thomas C., Icek Ajzen, and Daniel Hrubes. 2003. "Further tests of entreaties to avoid hypothetical bias in referendum contingent valuation." *Journal of Environmental Economics and Management* 46:353–361.

Brown, T.C., P.A. Champ, R.C. Bishop, and D.W. McCollum. 1996. "Which Response Format Reveals the Truth About Donations to a Public Good?" *Land Economics* 72(2):152–66.

Burton, Anthony C., Katherine S. Carson, Susan M. Chilton, and W. George Hutchinson. 2007. "Resolving Questions About Bias in Real and Hypothetical Referenda." *Environmental Resource Economics* 38: 513–525.

Carson, Richard, Theodore Groves, and John List. 2008. Probabilistic Influence and Supplemental Benefits A Field Test of the Two Key Assumptions Behind Using Stated Preferences. Draft: February. Bishop0000198_Carson-Groves-List 2008.pdf

Carson, R.T., W.M. Hanemann, R.C. Mitchell. 1986. *The Use of Simulated Political Markets to Value Public Goods.* Working Paper. Economics Department, University of California, San Diego.

Champ, Patricia A., and Thomas C. Brown. 1997. "A Comparison of Contingent and Actual Voting Behavior." Unpublished paper, Rocky Mountain Research Station, USDA Forest Service, Fort Collins, CO.

Cummings, Ronald G., Steven Elliott, Glenn W. Harrison, and James Murphy. 1997. "Are Hypothetical Referenda Incentive Compatible?" *The Journal of Political Economy* 105(3):609-621.

Dickie, M., A. Fisher, and S. Gerking. 1987. "Market transactions and hypothetical demand data: a comparative study." *J. Amer. Statist. Assoc.* 82(1): 69–75.

Duffield, J.W. and D.A. Patterson. 1992. "Field Testing Existence Values: An Instream Flow Trust Fund for Montana Rivers." Paper presented at the Allied Social Science Associations annual meeting. New Orleans, LA, January 4.

Johnston, Robert J. 2006. "Is Hypothetical Bias Universal? Validating Contingent Valuation Responses Using a Binding Public Referendum." *Journal of Environmental Economics and Management* 52:469-481.

Johnston, R.F., T.F. Weaver, L.A. Smith, S.K. Swallow. 1995. "Contingent valuation focus groups: insights from ethnographic interview techniques." *Agric. Resour. Econom. Rev.* 24(1): 56–69.

Kealy, M.J., M. Montgomery, and J.F. Dovidio. 1990. "Reliability and Predictive Validity of Contingent Values: Does the Nature of the Good Matter?" *Journal of Environmental Economics and Management* 19(3):244–263.

Knetsch, J.L. and R.K. Davis. 1966. "Comparisons of methods for recreation evaluation." In: A.V. Kneese, S.C. Smith (Eds.), *Water Research*, Resources for the Future Inc. Baltimore: John Hopkins Press. 125–142.

Landry, Craig E., and John A. List. 2007. "Using Ex Ante Approaches to Obtain Credible Signals for Value in Contingent Markets: Evidence from the Field." *American Journal Agricultural Economics* 89(2):420–429.

List, J.A and J.F. Shogren. 1998. "Calibration of the difference between actual and hypothetical valuations in a field experiment." *J. Econom. Behav. Organiz.* 37(2):193–206.

Loomis, J., M. Creel, and T. Park. 1991. "Comparing benefit estimates from travel cost and contingent valuation using confidence intervals for Hicksian welfare measures." *Appl. Econom.* 23(11): 1725–1731.

Murphy, J.J. and T.H. Stevens. 2004. "Contingent valuation, hypothetical bias, and experimental economics." *Agric. Resour. Econom. Rev.* 33(2): 182–192.

Schläpfer, Felix, Anna Roschewitz, and Nick Hanley. 2004. "Validation of Stated Preferences for Public Goods: A Comparison." *Ecological Economics* 51:1– 16

Seip, K. and J. Strand. 1992. "Willingness to Pay for Environmental Goods in Norway: A Contingent Valuation Study with Real Payment." *Environmental and Resource Economics* 2(1):91–106.

Shabman, L. and K. Stephenson. 1996. "Searching for the correct benefit estimate: empirical evidence for an alternative perspective." *Land Economics* 72(4): 433–449.

Taylor, Laura 0. 1998. "Incentive Compatible Referenda and the Valuation of Environmental Goods." *Agricultural and Resource Economics Review*, 27(2):132-39.

Taylor, Laura O., Michael McKee, Susan K. Laury, and Ronald G. Cummings. 2001. "Induced-Value Tests of the Referendum Voting Mechanism." *Economics Letters* 71:61–65

Vossler, Christian A. and Joe Kerkvliet. 2003. "A Criterion Validity Test of the Contingent Valuation Method: Comparing Hypothetical and Actual Voting Behavior for a Public Referendum." *Journal of Environmental Economics and Management* 45:631–649.

Vossler, Christian A., Joe Kerkvliet, Stephen Polasky, and Olesya Gainutdinova. 2003. Externally Validating Contingent Valuation: An Open-Space Survey and Referendum in Corvallis, Oregon. *Journal of Economic Behavior & Organization* 51:261–277.

Wang, Hua. 1997. "Treatment of 'Don't Know' Responses in Contingent Valuation Surveys: A Random Valuation Model." *Journal of Environmental Economics and Management.* 32: 219-232.

Appendix B

Divergence Between Base and Scope Survey

Divergence between the "base" and "scope" questionnaires in the Stratus study.

Yellow highlighting in the base survey denotes text not included in the scope survey. Red highlighting denotes numbers that are different.

The base scenario

Volume I, page 4-17 through page 4-34. Parts that are different form the "scope" scenario are highlighted:

"The ban and the other things being done will greatly reduce the amount of new phosphorus put onto land and in the river and lake in the future, but a lot of phosphorus that was spread on the land in the past will remain there. For many years, it will continue to wash into the river and lake when it rains.

"The purpose of this interview is to find out whether you think the State should or should not do something else as well. The excess phosphorus could be removed by putting alum on the land and in the water. I will tell you about what alum is and how it could be used to remove the excess phosphorus in a moment. After I tell you about the situation, I will ask you to vote on whether the state should or should not put alum on the land and in the water in order to return the river and lake to around 1960 conditions faster. Your vote will help state officials to decide whether to carry out the alum treatments.

"When alum is put into river or lake water that contains phosphorus, the alum attaches to the phosphorus to form harmless particles that fall to the bottom and blend into the dirt there. So if alum were put into the river and lake, the phosphorus there could no longer help algae to grow and there would then be a lot less algae in the water.
If alum is put on land, it attaches to phosphorus in the soil to form harmless particles. When these particles wash into rivers and lakes, the particles sink to the bottom and do not help algae to grow.
So to reduce algae in the river and lake, alum could be spread on the land and on the water.

"Here's how the alum treatments could be done.
The Army Corps of Engineers operates the lake, and they would work with the Oklahoma Department of Environmental Quality to spread the alum.
Crews of people would be hired and trained to use trucks to put alum on the land.
Specially designed boats would spread alum on the lake.
Alum would also remove phosphorus from river water flowing into Oklahoma from Arkansas. Dispensers would be put near the border to spread alum on the water when sensors find lots of phosphorus in it.

"For more than 35 years, alum has been used successfully and safely to remove phosphorus and reduce algae in many states, such as Colorado, Texas, Missouri, South Dakota, Florida, Wisconsin, and Washington. Those states had some rivers and lakes with lots of algae like the Illinois River and Tenkiller Lake. Experiences in those states have convinced scientists that alum does not harm fish or other things living in water, and that alum treatments here in Oklahoma could safely

return the river and lake to what they were like in around 1960.
Putting alum on the land and in the water would have some undesirable effects. The alum would be a white powder on the land surface until rains carry it down into the soil. After alum is put into the river and lake, it would make the water cloudy for a few hours until it settles to the bottom. And if anyone were to drink the lake water in the first hour, it might taste bitter.
Alum treatments would be needed for 5 years to remove all the excess phosphorus now on the land and in the water.

"(**POINT TO ROW 1**) A court-ordered ban would stop spreading of poultry litter near the river and lake in Oklahoma and Arkansas. This will occur even if alum treatments are not done.
(**POINT TO ROW 2**) Alum could be spread on land from trucks.
(**POINT TO ROW 3**) Alum could be spread on the lake from boats.
(**POINT TO ROW 4**) Alum could be sprayed in river water flowing into Oklahoma from Arkansas.
(**POINT TO ROW 5**) Alum treatments would need to be done for 5 years to remove all the excess phosphorus.

"As a result of alum treatments, the river would be back to what it was like in around 1960 (**POINT TO 1960**) about 10 years from now (**POINT TO 10 YEARS**). And the lake would be back to what it was like in around 1960 (**POINT TO 1960**) about [redacted] years from now (**POINT TO [redacted]**). Water in the river and lake would then be clear nearly all the time, and there would be little algae in the water and on the bottom. There would then be plenty of oxygen in the water. Species of fish, insects, small animals, and small plants that used to be common would slowly increase in numbers, replacing those that live in water with lots of algae. There would be fewer of some species, such as largemouth bass.

"The river and lake will go back to what they were like in around 1960 without alum, but it will take longer.
Scientists say that if spreading of poultry litter is banned, natural processes will allow the river and lake to gradually return to what they were like in around 1960, even with no alum treatments.
Rain would slowly wash the phosphorus into the river and lake for many years. Each year, a little less phosphorus would be washed into the river and lake. Because the river flows into the lake, the phosphorus in the river would be washed into the lake and would be kept there by the dam. The phosphorus would sink to the bottom of the lake and would slowly be covered by dirt, which would eventually seal it off, so that it could not help algae to grow.

"Without alum treatments, it will take about 50 years (**POINT**) for the river to get back to what it was like in around 1960 (**POINT TO 1960**) instead of about 10 years (**POINT**). That is about 40 years longer. It will take the lake about 60 years (**POINT**) to get back to what it was like in around 1960 (**POINT TO 1960**) instead of about [redacted] years (**POINT**). That is also about [redacted] years longer.

"If a court bans spreading of poultry litter, the industry will have to safely get rid

of all the litter they produce from now on. The industry will have to pay for this, and the river and lake will naturally return to what they were like in around 1960. If the people of Oklahoma want this to happen [redacted] years sooner, there will be an additional cost for the alum treatments. Oklahoma taxpayers will have to pay some of this cost because many chicken and turkey farms have gone out of business over the years. In addition, many other Oklahomans contributed to the excess phosphorus through sewage and their use of fertilizer.
We are interviewing people in Oklahoma to ask them to vote on whether the state should or should not put alum on the land and in the water. Your vote today will affect whether or not alum treatments are done.

"The state does not want to start the program unless it has all the funds needed to buy the equipment, hire and train the staff, and complete the 5 years of alum treatments. To pay for this, Oklahoma taxpayers would pay a one-time tax added to their state income tax bill next year. The cost to your household would be $(**BIDAMNT**). The money would go into a special trust fund that can be used only for alum treatments. This is the only payment that would be required.

"Voting for the program means (**PAUSE**) that it is worth it to you (**PAUSE**) for your household to pay the additional one-time tax of $ (**BIDAMT**) (**PAUSE**) to return the Illinois River, Flint Creek, Barren Fork Creek, the smaller creeks flowing into them, and Tenkiller Lake to what they were like in around 1960 [redacted] years sooner.

"(**POINT**) Natural processes will return the river and lake to what they were like in around 1960 in [redacted] 60 years without alum treatments.

"**Q25.** After spreading of litter is banned, how serious did you think the effects of algae in the river would be if no alum treatments are done? Not serious at all, slightly serious, moderately serious, very serious, or extremely serious?

"**Q30.** Now let's turn to the lake. I told you it would take about 60 years for the lake to return to what it was like in around 1960 without alum treatments. When you decided how to vote, did you think that it would take about 60 years, or did you think it would take less time or more time?

"**Q31.** When you decided how to vote, how well did you think that alum treatments would work at reducing algae in the water? Not well at all, slightly well, moderately well, very well, or extremely well?

"**Q33.** When you decided how to vote, did you think that the extra tax money would be used for alum treatments to reduce algae in only Tenkiller Lake and the Illinois River and creeks flowing into it, or did you think some of this money would be used clean up other rivers and lakes in Oklahoma as well?

"Q34. When you decided how to vote, did you think that if the alum treatments are done successfully for the Illinois River and Tenkiller Lake, this would or would not increase the chances that other rivers and lakes in Oklahoma would get alum treatments later?"

The scope scenario

Vol. I, page 4-17 through page 4-34:
"*The ban and the other things being done will greatly reduce the amount of new phosphorus put onto land and in the river and lake in the future. The excess phosphorus will quickly wash out of the river, but much of the phosphorus that's in the lake now will remain there.*

"*The purpose of this interview is to find out whether you think the State should or should not do something else in the lake. The excess phosphorus in the lake could be removed by putting alum in the water. I will tell you about what alum is and how it could be used to remove the excess phosphorus in the lake in a moment. After I tell you about the situation, I will ask you to vote on whether the state should or should not put alum in the lake in order to return the lake to around 1960 conditions somewhat faster. Your vote will help state officials to decide whether to carry out the alum treatments.*

"*When alum is put into lake water that contains phosphorus, the alum attaches to the phosphorus to form harmless particles that fall to the bottom and blend into the dirt there. So if alum were put into the lake, the phosphorus there could no longer help algae to grow and there would then be a lot less algae in the water. Alum treatments will not be needed for the river. The natural flow of water in the river will remove the excess phosphorus there. After the ban is in place, the river will naturally return to what it was like in around 1960 in [redacted] years. Phosphorus will remain in the lake much longer because the lake is large and the water moves through it very slowly.*

"Here's how the alum treatments would be done.
The Army Corps of Engineers operates the lake, and they would work with the Oklahoma Department of Environmental Quality to spread the alum.
Specially designed boats would spread alum on the lake.

"For more than 35 years, alum has been used successfully and safely to remove phosphorus and reduce algae in lakes in many states, including Colorado, Texas, Missouri, South Dakota, Florida, Wisconsin, and Washington. Those states had some lakes with lots of algae like Tenkiller Lake. Experiences in those states have convinced scientists that alum does not harm fish or other things living in lakes, and that alum treatments here in Oklahoma could safely return the lake to what it was like in around 1960.
Putting alum in the lake would have some undesirable effects. After alum is put into the lake, it would make the water cloudy for a few hours until it settles to the bottom. And if anyone were to drink the lake water in the first hour, it might taste bitter.
Alum treatments would be needed for 5 years to remove all the excess phosphorus in the lake.

"***(POINT TO ROW 1)*** *A court-ordered ban would stop spreading of poultry litter near the river and lake in Oklahoma and Arkansas. This will occur even if alum treatments are not done.*

*(**POINT TO ROW 2**) Alum could be spread on the lake from boats.*
*(**POINT TO ROW 3**) Alum treatments would need to be done for 5 years to remove all the excess phosphorus from the lake.*

*"As a result of alum treatments, the lake would be back to what it was like in around 1960 (**POINT TO 1960**) about [redacted] years from now (**POINT TO [redacted] YEARS**). Water in the lake would then be clear nearly all the time, and there would be little algae in the water and on the bottom. There would then be plenty of oxygen in the water. Species of fish, insects, small animals, and small plants that used to be common would slowly increase in numbers, replacing those that live in water with lots of algae. There would be fewer of some species, such as largemouth bass.*

"Scientists say that if spreading of poultry litter is banned, natural processes will gradually return the lake to what it was like in around 1960, even with no alum treatments, but it will take somewhat longer.
The phosphorus remaining in the lake would sink to the bottom and would slowly be covered by dirt, which would eventually seal it off, so that it could not help algae to grow.

*"Without alum treatments, it will take the lake about 60 years (**POINT**) to get back to what it was like in around 1960 (**POINT TO 1960**) instead of about [redacted] years (**POINT**). That is about [redacted] years longer.*

"If a court bans spreading of poultry litter, the industry will have to safely get rid of all the litter they produce from now on. The industry will have to pay for this. The river will naturally return to what it was like in around 1960 in [redacted] years, and the lake will naturally return to what it was like in around 1960 in [redacted] years. If the people of Oklahoma want the lake to return to what it was like in around 1960 in 50 years rather than 60 years, there will be an additional cost for the alum treatments. Oklahoma taxpayers will have to pay some of this cost because many chicken and turkey farms have gone out of business over the years. In addition, many other Oklahomans contributed to the excess phosphorus in the lake through sewage and their use of fertilizer.
We are interviewing people in Oklahoma to ask them to vote on whether the state should or should not put alum in the lake. Your vote today will affect whether or not alum treatments are done.

"*The state does not want to start the program unless it has all the funds needed to buy the equipment, hire and train the staff, and complete the 5 years of alum treatments to the lake. To pay for this, Oklahoma taxpayers would pay a one-time tax added to their state income tax bill next year The cost to your household would be $_(BIDAMNT). The money would go into a special trust fund that can be used only for alum treatments. This is the only payment that would be required.*

*"Voting for the program means (**PAUSE**) that it is worth it to you (**PAUSE**) for your household to pay the additional one-time tax of $ (**BIDAMT**) (**PAUSE**) to return Tenkiller Lake to what it was like in around 1960 in [redacted] years rather than [redacted] years.*

*"(**POINT**) Natural processes will return the lake to what it was like in around 1960 in 60 years without alum treatments.*

*"**Q25**. After spreading of litter is banned, how serious did you think the effects of algae in the river would be? Not serious at all, slightly serious, moderately serious, very serious, or extremely serious?*

*"**Q30.** I told you it would take about 60 years for the lake to return to what it was like in around 1960 without alum treatments. When you decided how to vote, did you think that it would take about 60 years, or did you think it would take less time or more time?*

*"**Q31.** When you decided how to vote, how well did you think that alum treatments would work at reducing algae in the lake? Not well at all, slightly well, moderately well, very well, or extremely well?*

*"**Q34**. When you decided how to vote, did you think that if the alum treatments are done successfully for Tenkiller Lake, this would or would not increase the chances that other lakes in Oklahoma would get alum treatments later?"*

Appendix C
Analysis of Subgroup Respondents

Figure C.1

		ABERS	95% Lower	95% Upper	Turnbull	95% Lower	95% Upper
	All	$184.55	$165.72	$203.38	$176.78	$160.09	$193.43
More/ same/ less spending on pollution	Less, Same	$119.96	-	-	$119.96	$98.35	$141.57
Visited IL River ever	No	$163.53	-	-	$154.21	$125.12	$183.31
	Yes	$202.89	-	-	$202.89	$180.68	$225.09
Visited TKL Lake ever	No	$149.89	-	-	$142.08	$114.13	$170.04
	Yes	$216.83	-	-	$135.00	$126.28	$143.72
Aware of scenic river status	No	$175.10	$151.32	$198.88	$175.10	$155.96	$194.24
	Yes	$204.11	-	-	$194.95	$160.05	$229.85
Heard of reason for phosphorous	No	$186.56	$161.66	$211.46	$186.56	$166.56	$206.56
	Yes	$180.69	-	-	$171.92	$139.09	$204.74
Effectiveness of Alum	Not, Slight, Moderate	$100.88	-	-	$100.88	$81.08	$120.68
Active User of IL River		$120.54	-	-	$120.54	$90.67	$150.41
Active User of TKL Lake		$120.65	-	-			
Passive User of TKL Lake		$85.09	-	-	$85.09	$62.59	$107.59
Effectiveness of Alum	Very, Extreme	$244.77	$220.08	$269.47			
Active User of IL River		$262.99	-	-	$257.53	$221.18	$293.88
Passive User of IL River		$223.47	-	-	$209.28	$168.69	$249.86
Tax goes to treat other lakes and rivers	No	$198.66	-	-	$198.66	$173.00	$224.32
	Yes	$179.43	$159.01	$199.85	$145.83	$173.13	$200.44
Plan to leave OK in next year	No	$178.93	$157.98	$199.89	$170.49	$147.53	$193.44
Paid taxes in 2007	Yes	$185.61	$167.14	$204.08	$180.87	$157.98	$199.89
Difficulty paying tax	Extreme, Very	$101.42	-	-	$94.79	$70.74	$118.84
	Slight, Not	$244.77	$220.08	$269.47	$235.29	$207.89	$262.68
Quartiles of HH income before taxes	y>$150,000	$189.02	-	-	$172.97	$104.26	$241.67
	$50,000<y<$150,000				$176.72	$148.34	$204.10
	$22,750<y<$50,000	$184.28	-	-	$176.80	$133.73	$219.87
	0<y<$22,750	$187.67	-	-			

Table C.1: Comparison of WTP for Active v. Passive Users of IL River

Q14: Have you ever visited the IL River?

	No (Passive Users)			Yes (Active Users)		
	WTP	Lower 95%	Upper 95%	WTP	Lower 95%	Upper 95%
ABERS	$163.53			$202.89		
TRNBL	$154.21	$125.12	$183.31	$202.89	$180.68	$225.09
	Pr(Yes|Bid)	Lower 95%	Upper 95%	Pr(Yes|Bid)	Lower 95%	Upper 95%
$10	0.75	0.75	0.75	0.85	0.85	0.85
$45	0.65	0.65	0.65	0.76	0.76	0.76
$80	0.57	0.57	0.57	0.64	0.64	0.64
$125	0.62	0.62	0.62	0.61	0.61	0.61
$205	0.39	0.39	0.39	0.48	0.48	0.48
$405	0.27	0.27	0.27	0.40	0.40	0.40

Table C.2: Comparison of WTP for Differences in Perceived Effectiveness

(Q31: How effective do you think the alum treatment will be?)

	Not, Slightly, Moderately			Very, Extremely		
	WTP	Lower 95%	Upper 95%	WTP	Lower 95%	Upper 95%
ABERS	100.88			244.77	220.08	269.47
TRNBL	100.88	81.08	120.68	235.29	207.89	262.68
	Pr(Yes|Bid)	Lower 95%	Upper 95%	Pr(Yes|Bid)	Lower 95%	Upper 95%
$ 10	0.65	0.52	0.52	0.65	0.52	0.52
$ 45	0.49	0.34	0.34	0.49	0.34	0.34
$ 80	0.44	0.59	0.59	0.44	0.59	0.59
$ 125	0.36	0.12	0.12	0.36	0.12	0.12
$ 205	0.24	0.00	0.00	0.24	0.00	0.00
$ 405	0.13	0.00	0.00	0.13	0.00	0.00

Table C.3: Comparison of WTP for Different Views Regarding State Spending

Q7e: Should the state spend less, same, or more on pollution?

	Less, Same			More		
	WTP	**Lower 95%**	**Upper 95%**	**WTP**	**Lower 95%**	**Upper 95%**
ABERS	119.96					
TRNBL	119.96	98.35	141.57			
	Pr(Yes)	Lower 95%	Upper 95%	Pr(Yes)	Lower 95%	Upper 95%
$ 10	0.67	0.67	0.67	0.67	0.67	0.67
$ 45	0.60	0.60	0.60	0.58	0.58	0.58
$ 80	0.57	0.57	0.57	0.56	0.56	0.56
$ 125	0.51	0.51	0.51	0.52	0.52	0.52
$ 205	0.35	0.35	0.35	0.37	0.37	0.37
$ 405	0.11	0.11	0.11	0.10	0.10	0.10

Table C.4: WTP and Confidence Intervals for Income Quintiles

	ABERS	**Turnbull**	**Lower 95%**	**Upper 95%**
$64,000<y<$600,001	$158.78	$139.33	$67.47	$211.18
$43,000<y<$64,000	$192.85	$157.10	$141.15	$173.04
$27,000<y<$43,000	$188.38	$188.38	$151.29	$225.47
$15,000<y<$27,000	$198.91	$195.16	$155.28	$235.04
$0<y<$15,000	$183.72	$160.09	$125.70	$194.48

Table C.5: WTP and Confidence Intervals for Income Sextiles

	ABERS	**Turnbull**	**Lower 95%**	**Upper 95%**
$70,000<y<$60,000	$145.93	$121.11	$50.84	$191.38
$50,000<y<$70,000	$216.26	$72.07	$65.01	$79.13
$33,000<y<$50,000	$178.89	$166.00	$122.54	$209.47
$23,000<y<$33,000	$186.87	$186.87	$151.83	$221.91
$13,000<y<$23,000	$172.08	$161.20	$100.95	$221.44
$0<y<$13,000	$202.14	$192.04	$135.39	$248.69

Appendix D

WTP Estimates with Recoded Data

Table D.1: ABERS WTP Estimates for Original and Recoded Data

Dataset	WTP Estimate	% Decline
Stratus Original Estimate	**$184.55**	**0.0%**
Recode: Alum Without Ban	$153.39	-16.9%
Recode: Certainty	$131.62	-28.7%
Recode: Faster - River	$144.62	-21.6%
Recode: Faster - Lake	$141.92	-23.1%
Recode: Pay Other River/Lake	$103.15	-44.1%
Recode: Income Tax	$116.47	-36.9%
Combined Recodes	**$37.98**	**-79.4%**

Table D.2: ABERS WTP Estimates for Passive and Active Users of the Illinois River

	Passive IL River		Active IL River	
Dataset	ABERS WTP	% Decline	ABERS WTP	% Decline
Stratus Original	**$163.53**	**0.0%**	**$202.89**	**0.0%**
Recode: Alum Without Ban	$129.42	-20.9%	$174.54	-14.0%
Recode: Certainty	$117.20	-28.3%	$144.16	-28.9%
Recode: Faster - River	$115.00	-29.7%	$115.56	-43.0%
Recode: Faster - Lake	$114.33	-30.1%	$165.71	-18.3%
Recode: Pay Other River/Lake	$89.13	-45.5%	115.56	-43.0%
Recode: Income Tax	88.68	-45.8%	$139.52	-31.2%
Combined Recodes	**$29.50**	**-82.0%**	**18.8**	**-90.7%**

Table D.3: ABERS WTP Estimates for Passive and Active Users of Tenkiller Lake

	Passive Lake		Active Lake	
Dataset	ABERS WTP	% Decline	ABERS WTP	% Decline
Stratus Original	**$149.89**	**0.0%**	**$216.83**	**0.0%**
Recode: Alum Without Ban	$114.89	-23.4%	$191.34	-11.8%
Recode: Certainty	$108.75	-27.4%	$154.36	-28.8%
Recode: Faster - River	$109.35	-27.0%	$174.35	-19.6%
Recode: Faster - Lake	115.29	-23.1%	-	-
Recode: Pay Other River/Lake	$87.39	-41.7%	$119.30	-45.0%
Recode: Income Tax	$95.20	-36.5%	$137.40	-36.6%
Combined Recodes	**$32.55**	**-78.3%**	**$42.65**	**-80.3%**

Table D.4: ABERS WTP Estimates by respondents' belief in alum treatment effectiveness: Not, Slightly, Moderate versus Very or Extremely

Dataset	Not/Slightly/ Moderate	% Decline	Very/ Extremely	% Decline
Stratus Original	**$100.88**	**0.0%**	**$244.77**	**0.0%**
Recode: Alum Without Ban	$82.10	-18.6%	$204.38	-16.5%
Recode: Certainty	$46.92	-53.5%	$189.34	-22.6%
Recode: Faster - River	$71.06	-29.6%	$196.69	-19.6%
Recode: Faster - Lake	$67.86	-32.7%	$194.47	-20.5%
Recode: Pay Other River/Lake	$48.50	-51.9%	$142.43	-41.8%
Recode: Income Tax	$69.39	-31.2%	$149.34	-39.0%
Combined Recodes	**$10.74**	**-89.4%**	**$56.15**	**-77.1%**

Table D.5: ABERS WTP Estimates by Respondents' Belief that state should spend less, or the same on pollution (Q7e)

Dataset	WTP Estimate	% Decline
Stratus Original	**$119.96**	**0.0%**
Recode: Alum Without Ban	$106.64	-11.1%
Recode: Certainty	$86.84	-27.6%
Recode: Faster - River	$94.20	-21.5%
Recode: Faster - Lake	$90.79	-24.3%
Recode: Pay Other River/Lake	$65.82	-45.1%
Recode: Income Tax	$62.09	-48.2%
Combined Recodes	**$28.58**	**-76.2%**

Appendix E
Income Elasticities by Quartile, Quintile, and Sextile for Recoded Data

Table E.1: Income Elasticities by Income Quartile

	Quartile	Max Income	Income Elasticities
Stratus	1	$600,001	0.20
	2	$60,000	0.18
	3	$33,000	-0.40
	4	$18,000	
All revisions	1	$600,001	
	2	$60,000	
	3	$33,000	0.11
	4	$18,000	-0.09
Ban	1	$600,001	0.06
Ban	2	$60,000	-0.14
Ban	3	$33,000	-0.38
Ban	4	$18,000	0.09
Certainty	1	$600,001	0.25
Certainty	2	$60,000	0.21
Certainty	3	$33,000	0.22
Certainty	4	$18,000	0.12
Faster - River	1	$600,001	0.57
Faster - River	2	$60,000	0.96
Faster - River	3	$33,000	-0.84
Faster - River	4	$18,000	0.02
Faster - Lake	1	$600,001	0.16
Faster - Lake	2	$60,000	1.12
Faster - Lake	3	$33,000	-1.24
Faster - Lake	4	$18,000	0.52
Other River/Lake	1	$600,001	0.37
Other River/Lake	2	$60,000	-1.25
Other River/Lake	3	$33,000	-0.02
Other River/Lake	4	$18,000	0.08
Income Tax	1	$600,001	0.20
Income Tax	2	$60,000	-0.59
Income Tax	3	$33,000	-0.71
Income Tax	4	$18,000	0.19

Table E.2: Income Elasticities by Income Quintile

	Quartile	Max Income	Income
Stratus	1	$600,001	0.200
Stratus	2	$60,000	0.177
Stratus	3	$33,000	-0.397
Stratus	4	$18,000	
All revisions	1	$600,001	
All revisions	2	$60,000	
All revisions	3	$33,000	0.113
All revisions	4	$18,000	-0.091
Ban	1	$600,001	0.064
Ban	2	$60,000	-0.137
Ban	3	$33,000	-0.384
Ban	4	$18,000	0.088
Certainty	1	$600,001	0.254
Certainty	2	$60,000	0.214
Certainty	3	$33,000	0.221
Certainty	4	$18,000	0.122
Faster - River	1	$600,001	0.572
Faster - River	2	$60,000	0.960
Faster - River	3	$33,000	-0.836
Faster - River	4	$18,000	0.024
Faster - Lake	1	$600,001	0.159
Faster - Lake	2	$60,000	1.122
Faster - Lake	3	$33,000	-1.236
Faster - Lake	4	$18,000	0.517
Other River/Lake	1	$600,001	0.375
Other River/Lake	2	$60,000	-1.247
Other River/Lake	3	$33,000	-0.022
Other River/Lake	4	$18,000	0.080
Income Tax	1	$600,001	0.196
Income Tax	2	$60,000	-0.588
Income Tax	3	$33,000	-0.711
Income Tax	4	$18,000	0.190

Table E.3: Income Elasticities by Income Sextile

	Sextile	Max Income	Income
Stratus	1	$600,001	-0.09
Stratus	2	$70,000	
Stratus	3	$50,000	1.51
Stratus	4	$33,000	-0.59
Stratus	5	$23,000	-0.70
Stratus	6	$13,000	0.16
All revisions		-$25	
All revisions		$0	
Ban	1	$600,001	-0.21
Ban	2	$70,000	0.68
Ban	3	$50,000	1.59
Ban	4	$33,000	-0.51
Ban	5	$23,000	-0.92
Ban	6	$13,000	0.06
Certainty	1	$600,001	0.00
Certainty	2	$70,000	1.06
Certainty	3	$50,000	4.42
Certainty	4	$33,000	0.55
Certainty	5	$23,000	-3.78
Certainty	6	$13,000	-0.05
Faster – River	1	$600,001	0.05
Faster – River	2	$70,000	-0.38
Faster – River	3	$50,000	4.65
Faster – River	4	$33,000	-1.05
Faster – River	5	$23,000	-2.82
Faster – River	6	$13,000	-0.27
Faster – Lake	1	$600,001	-0.07
Faster – Lake	2	$70,000	1.34
Faster – Lake	3	$50,000	3.67
Faster – Lake	4	$33,000	-1.70
Faster - Lake	5	$23,000	-2.14
Faster - Lake	6	$13,000	-0.37
Other River/Lake	1	$600,001	-2.23
Other River/Lake	2	$70,000	0.00
Other River/Lake	3	$50,000	7.26
Other River/Lake	4	$33,000	-0.57
Other River/Lake	5	$23,000	-2.41
Other River/Lake	6	$13,000	-0.06
Income Tax	1	$600,001	-5.94
Income Tax	2	$70,000	0.00
Income Tax	3	$50,000	3.02
Income Tax	4	$33,000	-0.69
Income Tax	5	$23,000	0.12
Income Tax	6	$13,000	-0.99